Environmental Science
Water Research & Technology

Stormwater subsurface gravel wetland hydraulics, phosphorus retention, and chloride dynamics in cold climates

Journal:	*Environmental Science: Water Research & Technology*
Manuscript ID	EW-ART-01-2023-000062.R1
Article Type:	Paper

Water Impact Statement

The stormwater treatment performance of an increasingly popular horizontal subsurface-flow gravel wetland design in the northeastern United States falls far short of expectations. Urban field installations studied were characterized by negligible phosphorus retention. A laboratory experiment demonstrated that dissolved phosphorus export is likely driven by phosphorus leaching from engineered soil. We recommend a testing protocol to guide soil selection and improve performance.

ARTICLE

Stormwater subsurface gravel wetland hydraulics, phosphorus retention, and chloride dynamics in cold climates

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Eric D. Roy,*[a,b,c] Andres O. Torizzo,[d] Marcos L. Kubow,[a] Nisha C. Nadkarni,[d] Thomas M. Adler,[d] Madeline F. Yandow,[d] Finn A. Bondeson,[b] Adrian R.H. Wiegman,[ac] and Donna M. Rizzo[b,c]

Subsurface gravel wetlands (SGW) are water treatment practices that use a saturated layer of gravel, sometimes below a vegetated soil layer, to filter urban stormwater and remove pollutants during horizontal flow. In recent years, the implementation of SGWs has proliferated among municipalities in the Northeast United States to meet phosphorus (P) control requirements. However, stormwater SGW performance is not well researched, creating knowledge gaps related to P removal performance and the effects of road salt on performance in cold climates. Here, we used field monitoring of two SGWs over two years and a complementary series of laboratory studies to examine SGW hydraulics, P retention, and chloride dynamics. Field results indicated reductions in peak flows and flow volumes, net P export for approximately half of the total storms monitored across a wide range of influent P loads, and chloride load reductions at one site. Lab results showed that both engineered soils and native soils are unlikely to restrict hydraulic conductivity to the degree desired. Furthermore, two out of three engineered soils tested, including the one used at the field sites, can release substantial P post-installation, while gravels have limited ability to sorb dissolved P. Neither soils nor gravels substantially influenced chloride concentrations in the lab experiments, and results illustrated two potential responses of wetland vegetation to chloride exposure in SGWs as well as chloride assimilation by vegetation. We recommend a testing protocol based on Mehlich-3 P saturation ratio to guide soil selection and reduce dissolved P leaching risk.

1. Introduction

Subsurface gravel wetlands (SGW) are water treatment practices that utilize horizontal flow through a saturated bed of gravel, sometimes below a vegetated soil layer, to filter stormwater runoff and remove pollutants through a combination of physical filtration, adsorption, biological uptake, and microbial transformation[1]. Water level is controlled by an outlet structure to retain a permanent subsurface pool, with ephemeral storage above the gravel to provide reduction of peak flows in addition to pollutant removal[2]. SGWs are distinct from bioretention in that they (a) host subsurface horizontal flow through a saturated gravel layer rather than vertical infiltration through bioretention media and gravel and (b) are never designed to facilitate infiltration into local soil[2,3]. SGWs are becoming increasingly popular tools for stormwater treatment in the Northeast United States, including in the State of Vermont, where they are considered a preferred ("Tier 2") practice for stormwater treatment when Tier 1 infiltration practices are infeasible and are expected to offer substantial (60-80%) reduction in total phosphorus (P) loads[3]. This is a strong driver of implementation given ongoing efforts to meet P load reductions required by, for example, the U.S. Environmental Protection Agency's Total Maximum Daily Load for P in the Vermont portion of the binational Lake Champlain Basin[4].

[a] Rubenstein School of Environment & Natural Resources, University of Vermont, Burlington, VT, USA 05405. Email: eroy4@uvm.edu
[b] Department of Civil & Environmental Engineering, University of Vermont, Burlington, VT, USA 05405
[c] Gund Institute for Environment, University of Vermont, Burlington, VT, USA 05405
[d] Watershed Consulting Associates LLC, Burlington, VT, USA 05401

While there is a rich literature focused on decades of SGW application in the context of wastewater treatment[5,6], SGWs treating urban stormwater, which is typically characterized by more dilute nutrient concentrations and more variable inflows, have been far less studied. Some evidence supports that SGWs can achieve >50% total P (TP) removal for stormwater[7-9]. However, others have documented negligible or even negative TP retention, especially where stormwater influent concentrations are low[10]. To date, urban SGWs have been excluded from the International Stormwater Best Management Practice (BMP) Database and are underrepresented in terms of data collected[11].

Challenges with premature adoption or flawed design of stormwater controls intended for urban stormwater P removal have been documented for other BMPs. For example, stormwater ponds were widely used for decades to control peak discharge and presumably water quality across North America and Europe[8,12,13]. However, in recent years, accumulating evidence has called into question the capacity of ponds to effectively mitigate the negative impacts of urban development on waterways simply by controlling peak discharge, and internal release of dissolved P from pond sediments has been linked to lower than expected total P retention in some cases[14-18]. Another example is stormwater bioretention systems, which have often been found to provide poor P retention[11] due to insufficient P sorption capacity[19,20] and leaching of P from compost used to support plant growth[21,22]. Therefore, poorly characterized alternatives to these popular BMPs, such as SGWs, should be considered with caution, and more evidence is needed to inform design.

Emerging concerns about the impact of chloride (Cl^-) on natural waters and roadside soils in cold climates raises questions regarding

the impact of road salt-laden runoff entering vegetated treatment practices such as SGWs[23]. In areas where salt application is significant, such as highways and state roadways, chloride loading to stormwater practices may change the hydraulic conductivity and/or vegetation health and survival – rendering practice performance different than what models predict[24-26]. Chloride concentrations in streams of Maryland, New York, and New Hampshire have reached as high as 5,000 mg L^{-1} in winter and elevated chloride concentrations can persist into the summer[27]. While the majority of chloride may pass directly through the soil and gravel layers of SGWs[28], chloride's solubility also makes it available to plants. Chloride tolerance varies by species, and some recommend planting "halophytic" or salt tolerant plants, as chloride is toxic to plants in excess[29-31].

In this study, we monitored two newly constructed SGWs in Vermont, USA and utilized complementary laboratory experiments to evaluate SGW performance and clarify opportunities to improve the design of SGWs treating urban stormwater in cold climates. We focused on both phosphorus and chloride due to ongoing widespread efforts to mitigate loading of these two constituents to freshwater ecosystems in cold climates. Our study objectives were to:

(1) Determine whether two recently installed SGWs in urban settings are performing as expected for flow attenuation and P retention.

(2) Determine how design variables, including material selection for SGW soils and gravel, affect P retention in a controlled laboratory experiment, and provide recommendations for future material screening.

(3) Determine how chloride is moving through and being stored within SGWs in the field.

(4) Determine (a) whether SGW material selection affects chloride transport and (b) how chloride affects plant species commonly used in local SGWs, using controlled laboratory experiments.

2. Materials and methods

2.1. Description of field sites

Two SGW field sites were monitored during 2020 and 2021. Both sites were designed based on guidance now being used in multiple states of the Northeast USA[2,3]. The first SGW site, "Fairview", is located in Essex Junction, Vermont, USA (44.49938, -73.09793) and has a drainage area of 9.4 ha, 1.6 ha of which are impervious, 5.3 ha are grassed, and 2.5 ha are wooded. The site consists of two bays, a pre-treatment forebay (85 m^2) and a treatment bay (380 m^2) (Figure 1a). Two main inlets feed into the forebay. The forebay is significantly elevated from the treatment bay, to which it is connected via a pipe and a rock lined emergency spillway for high flow events that exceed the capacity of the forebay. Pre-treated water leaves the forebay via an outlet structure that discharges to a 30.5 cm solid high-density polyethylene (HDPE) pipe that delivers runoff to 30.5 cm perforated underdrains within the subsurface gravel. The pre-treated water then flows horizontally through the saturated gravel media layer. There are also two inlets feeding directly into the treatment bay. The first is a 30.5 cm culvert bringing drainage from an adjacent road. It has a drainage area of 0.085 ha, 0.004 ha of which is grassed, and the rest is impervious area of the road surface. The other inlet is a 38.1 cm culvert bringing drainage from two roads. It has a drainage area of 0.61 ha, 0.22 ha of which is impervious, 0.21 ha is grassed, and 0.17 ha is wooded. A series of 30.5 cm perforated PVC underdrains are located on the southern side of the treatment bay and collect the treated water and convey it to the outlet structure. Treated water then enters the outlet structure and is discharged via a 61 cm HDPE pipe that connects to existing drainage structures that pass under a road and enters a small stream. Construction of the Fairview SGW was completed in the fall of 2019 and seeded with a wetland seed mix. Wetland plants were beginning to establish root structures by the end of the 2020 growing season (Figure S1, Supplementary Materials).

The second SGW site, "Kennedy" is located in South Burlington, Vermont, USA (44.45365, -73.16999) and has a drainage area of 3.47 ha, 2.47 ha of which are impervious, and 1.00 ha are grassed. This site consists of three bays - one 38 m^2 forebay and two treatment bays (195 m^2 and 143 m^2) - separated by berms (Figure 1b). Influent discharges into the forebay via a drainage structure that collects stormwater from two sources: a stone-lined swale that runs along the adjacent road and a storm drain crossing under the adjacent road. Influent enters the pre-treatment forebay via a 61 cm HDPE solid pipe. Influent overtops the berm and flows into the first treatment bay. During moderate to larger storms, influent can overtop the berm and access the first treatment bay via a spillway. At the near end of the first treatment bay, influent enters subsurface perforated pipes, then moves horizontally through the subsurface gravel layer before passage to the second treatment bay via a 61 cm HDPE pipe. At the near end of the second treatment bay, water enters the subsurface perforated pipes for passage through the gravel medium to the far end of the treatment bay. Effluent departs the final treatment bay through a solid pipe and is discharged into a natural wetland draining to a stream. A stone-lined emergency spillway is located along the western edge of the second treatment bay. Construction of the Kennedy SGW was completed in the fall of 2019 and seeded with a wetland seed mix. Vegetation establishment was slower for the Kennedy system compared to Fairview, with greater biomass present in 2021 compared to 2020 (Figure S2, Supplemental Materials).

At both sites, the treatment bays consisted of a 20.3 cm surface layer of an engineered soil material obtained from a local compost and soil supplier (labelled "em1" in the laboratory study described below). Below that, an 8-10 cm layer pea gravel was used to help restrict movement of fines into the gravel layer below, which was 61 cm deep and consisted of quartzite gravel (3.8 cm at Fairview, 1.9 cm at Kennedy). Soils were seeded with a wetland seed mix (see Supplementary Materials for details). Kennedy included a geotextile liner beneath the gravel layer, whereas Fairview did not.

2.2. Field monitoring of stormwater flows and water quality

For field monitoring, a Teledyne 6712 ISCO Portable Sampler was placed at each inlet and outlet riser, along with a specific conductance logger (In-Situ Aqua TROLL 600 Data Logger at inlet, In-Situ Aqua TROLL 200 Data Logger at outlet). An Onset HOBO Fresh Water Conductivity Data Logger was placed in a riser pipe in the centre of each treatment bay. Probes and strainers for the samplers/loggers were placed at the bottom of each riser. For the Fairview site, Thel-mar weirs were installed at the ends of two inlet culverts entering the treatment bay directly. HOBO water level loggers (Onset HOBO MX2001-04-SS)

measured water depth and temperature during each storm event at 5 min intervals at the inlets, outlets, and midpoints of each SGW, and were used to crosscheck the depth recorded by the ISCO flow modules. Water quality sampling was initiated when a water level rise ≥3 cm was detected at each inlet and outlet culvert and the ISCOs continued to collect 100 ml samples over a 24-hour period to produce the composite sample on a flow-weighted interval, for a total volume of 18 L. For larger storms, the ISCO samplers operated for a 48-hour period. The anticipated flow volume was based upon predicted rainfall and modelled flow volume for each outlet, with the sample program adjusted for each event.

During water sampling, the ISCO 6712 units were packed with ice packs. All containers in each ISCO unit were washed and tripled rinsed with sample water prior to sampling. Composite samples were retrieved after each sampling event, mixed, and subsampled into plastic bottles triple rinsed with sample water for total suspended solids (TSS) (SM 2540 D-11) and chloride (Cl^-) (EPA 300.0) measurement and new glass vials for total phosphorus (TP) and total dissolved phosphorus (TDP) analyses (SM 20 4500 P-F). Field blanks were collected every third sampling event for one analyte and duplicates were taken every fourth sampling event for all of the analytes at one sampling location. Samples were retrieved within 24 hours of storm event conclusion, kept in a dark cooler with ice, and transported to Endyne Laboratory in Williston, VT, USA for analysis within 4 hrs of subsampling.

A handheld YSI Professional Plus Multiparameter Instrument was used to collect chloride measurements at each SGW inlet, outlet, and midpoint. After subsampling of ISCO-collected samples was complete, YSI measurements for chloride were also taken within the ISCO sample bottles in the 2021 sampling season.

In 2020, an ISCO 720 Submerged Probe and Flow Module was installed at the inlet and outlet risers at both Fairview and Kennedy. Manning's calculation was used for the inlet and outlet risers at both sites using the following variables measured in field: water level dataset (level per minute), pipe shape (round), pipe diameter, pipe slope, and pipe roughness. To reduce equipment challenges, in 2021, an ISCO 750 Area-Velocity Probe and Flow Module was installed at the inlet and outlet risers at both sites using the following variables measured in-field: water level dataset (level per minute), velocity dataset (velocity per minute), pipe shape (round), pipe diameter, and silt level. In addition, for both years, ISCO 730 Bubbler Flow Modules were connected to Thel-mar weirs at two inlet culverts to the Fairview SGW. At each culvert, a flow volume for each storm event was calculated using the following variables measured in field: water level (level per minute), weir type (V-notch), and weir size (90 degrees). In all cases, the resulting product was a total flow volume per storm event for each inlet and outlet.

The total calculated volume at the inlet or outlet was multiplied by the concentration of the pollutant during the course of the storm event. This concentration was determined using a composite of stormwater sampled at the inlet or outlet over the course of the storm event.

2.3. Field measurements of plant biomass, plant chloride, and soil chloride

Destructive harvesting of herbaceous biomass was conducted in October of 2020 for standing dead and fresh litter and in August of 2021 for live biomass. These harvests began with establishing randomly selected cross-sections at the inlet, midpoint, and outlet of the two SGWs. Along each cross section, plant biomass samples were collected from three 0.25 m^2 quadrants with stems clipped to within 1-2 cm of the sediment surface. Samples from the three quadrants along each cross-section were composited, homogenized, and placed in a cooler for transport to the lab. All sampling of plant biomass was conducted on dry days as samples preserve better when moisture is low. Samples from each cross-section were placed into large paper bags and set in a drying room at 60°C until constant weight was achieved (~1 week). Dry weight was recorded, samples were ground using a Wiley Mill, and then subsamples were sent to the University of Maine to be analysed for chloride via ion chromatography (EPA 9251 after $CaSO_4$ extraction).

Soil core sampling was performed on the eight-inch wetland soil layer the Fairview and Kennedy sites August 2021 and October of 2020, respectively. Sampling began with establishing randomly selected cross-sections at the inlet, midpoint, and outlet of the two SGWs. A modified 60 mL plastic syringe with serrated end was pressed into the soil layer until it was filled with approximately 50 mL of soil. Four syringes of soil were cored per cross-section, composited, homogenized, and placed in a cooler for transport to the lab. Samples were submitted to Endyne, Inc. in Williston, VT, USA for Cl^- (EPA 9056A) and electrical conductivity (mod. EPA 9050A) analyses.

2.4. Laboratory soil and gravel column studies

Soil materials tested included three engineered materials obtained directly from compost/soil facilities in the Northeast USA (em1, em2, em3), one of the engineered materials following ~1 y in place at one of the field sites (em1_f), and native soils from the two field sites (ns1, ns2). All soil materials were analysed for total solids, total carbon, volatile solids, total nitrogen, total phosphorus, Modified Morgan phosphorus, Mehlich-3 phosphorus, Mehlich-3 iron, Mehlich-3 aluminium, bulk density, pH, and electrical conductivity at the University of Maine Analytical Lab and Maine Soil testing Service[32]. For Modified Morgan phosphorus, air-dry, 2 mm sieved soil samples were extracted with Modified Morgan extractant (0.62 N NH_4OH + 1.25 N CH_3COOH) by shaking a soil-solution suspension for 15 minutes at a 1:5 (soil:solution) mass ratio followed by filtering to remove particles above 8 µm in size. Extracts from the Modified Morgan procedure were analysed for P by ICP-OES. For Mehlich-3 extractions, air-dry, 2 mm sieved soil samples were extracted with the Mehlich-3 solution (0.2 M CH_3COOH + 0.25 M NH_4NO_3 + 0.015 M NH_4F + 0.013 M HNO_3 + 0.001 M EDTA) by shaking a soil-solution suspension for 5 minutes at a 1:10 (soil mass in g : solution volume in mL) ratio, followed by filtering to remove particles above 2 µm. Extracts from the Mehlich-3 procedure were analysed for P, iron (Fe), and aluminium (Al) by ICP-OES. For oxalate extractions at the University of Vermont, air-dried sieved (< 2 mm) and ground soils (0.2 g) were extracted in the dark 1:50 (weight/volume) with a pH 3.0 acid ammonium oxalate solution following Courchesne and Turmel[33] and analysed for P, Fe, and Al by ICP-OES.

The phosphorus saturation ratio (PSR) was calculated using both Mehlich-3 (PSR_{M3}) and oxalate (PSR_{OX}) extracts as:

$$PSR = \frac{\left(\frac{P}{31}\right)}{\left(\frac{Fe}{56}\right) + \left(\frac{Al}{27}\right)}$$

where, P = P in extract in mg P kg^{-1} dry soil, Fe = Fe in extract in mg Fe kg^{-1} dry soil, and Al = Al in extract in mg Al kg^{-1} dry soil[34,35].

Each soil material was also tested for water extractable P (WEP) at UVM using methods adapted from Kleinman et al.[36], including a 1:100 solid:solution mass ratio. Hydraulic conductivity was measured for all soil materials at UVM using a saturated hydraulic conductivity (Ksat) meter from Meter Group™ (n = 4 per soil). Soils were loosely packed in the Ksat meter to imitate field condition at time of SGW construction.

Sequential soil and gravel column experiments were designed to simulate anticipated hydraulics within SGWs and isolate material effects on the dynamics of suspended solids, phosphorus, and chloride in stormwater (Figure 2). For the soil columns, a 20 cm uncompact layer of soil material was placed in a 41 cm tall PVC pipe (3.8 cm radius) with a perforated (0.16 cm holes) central PVC drainpipe (1.3 cm radius) double wrapped in matted geotextile. Synthetic stormwater (0.2 mg PO_4-P L^{-1}, 0.5 mg NO_3-N L^{-1}, 0.5 mg NH_4-N L^{-1}, and 650 mg NaCl-Cl L^{-1}) was initially added down the drainpipe while the outlet valve was closed. Once the column reached a 20 cm ponding depth above the soil, the synthetic stormwater was incubated for 1 hour, after which the column outlet valve was opened to allow water to move through the perforated pipe and drain into a collection bucket underneath the column at a rate of 3-6 cm^3 s^{-1}. At the same time, synthetic stormwater was added to the water column above the soil using a peristaltic pump to maintain 20 cm ponding depth. For each simulation, a total of 3 L of synthetic stormwater was loaded to the soil column. This resulted in contact times ranging from approximately 15 to 75 min in total within each simulation, in line with expectations that standing water of significant depth is not expected in SGWs except for large storms[2], the standard duration (60 min) of water-extractable P protocols[36], and the time needed to represent initial rapid P desorption observed for soils[37]. At the end of each simulation, collected effluent was mixed and a subsample was collected for water quality analyses. Six simulations were performed for each column, with triplicate columns for each material tested. Soluble reactive P (SRP) and Cl^- were measured for all six simulations per soil core, while TP and TSS were measured for simulations 1, 3, and 6. Two separate soil column experiments were performed, each testing three treatment materials and a control blank (i.e., columns without any soil substrate receiving the same synthetic stormwater as the treatments during each experiment, with the same 1 hr incubation time, draining, and water quality analyses). The first included three engineered muck materials, while the second included one engineered muck material after ~1 year of time in the field and two native soils (one from each field site). For the first experiment, all effluent from treatment columns not used for water quality sampling was composited and preserved at 4°C for ≤ 1 week for the subsequent gravel column testing.

Gravel materials tested included three locally available angular gravels (1.3 to 1.9 cm): granite, quartzite, and limestone. A composite of all remaining soil column effluent from the first soil column experiment was pumped vertically upwards through the gravel columns (n = 3 per gravel type, 3.8 cm radius, 61 cm height) to achieve a hydraulic retention time (HRT) of 24 hours (+/- 30 min) for 6 days. This HRT aligns with expected hydraulics in field SGWs for a 2.5 cm storm. Every 24 hours, 200 mL was sampled from the inflow and outflow of each gravel column. SRP and Cl^- were measured for each sample, while all six inflow samples and all six outflow samples from each column were composited to obtain a single composite inflow sample and a single composite outflow sample for TP and TSS measurements for each column. Triplicate control blank columns without gravel but having the same HRT were also tested at the same time as the treatments, with identical water quality analyses.

Methods used for water quality analysis in the column experiments were as follows. SRP samples were filtered (0.45 µm) and frozen for storage, then thawed for analysis where absorbance at a wavelength of 660 nm was read on a microplate reader (Synergy HT, BioTek, Winooski, VT USA) using the malachite green method[38-40]. TP samples were digested following the alkaline persulfate digestion[41], and analysed using colorimetric orthophosphate analysis at 880 nm on a Lachat QuickChem 8500 using the ascorbic acid method for molybdenum blue[42]. TSS content was quantified by filtering a measured volume of water through a pre-ashed and weighed glass fibre filter, dried at 105°C, and then again weighed[43]. Cl^- and pH were measured using an YSI Professional Plus Multiparameter Instrument.

For each soil column experiment, results for effluent SRP, TP, and TSS were compared across treatment groups using one-way Kruskal-Wallis tests followed by Dunn-Bonferonni post-hoc comparisons (α = 0.5) in RStudio. For soil column experiments where potential trends in concentrations were observed for some soils across successive simulation events, a linear model of the format $\log_{10}(x) \sim$ soil × simulation was also examined. For TSS, x+1 was used in place of x in this model. For gravel column experiments, the same approach taken with soil columns for statistical analysis was used for effluent SRP. For TP and TSS in gravel column effluents, Tukey HSD contrasts using $\log_{10}$ transformed concentrations were used to compare across treatments (α = 0.5).

2.5. Laboratory chloride bioassay

A greenhouse bioassay approach derived from methods presented in Powell et al.[44] was used to further investigate chloride effects on vegetation included in SGWs. Shallow Sedge (*Carex lurida*) and Broad-fruited Bur-reed (*Sparganium eurycarpum*), both included in a commonly used wetland seed mix for local green infrastructure projects, were grown under controlled greenhouse conditions to measure stress response at varying levels of Cl^- exposure. Plants were acquired as plugs from Vermont Wetland Plant Supply Company with plant heights ranging between 15 and 30 cm and transplanted into 10 cm diameter pots filled with a garden soil mix (20% silt, 15% sand, 10% peat moss, and 55% compost). During the bioassay, all pots were kept moist in 16 randomized block pattern trays (46 cm x 5 cm x 9 cm) and sub-irrigated from the tray. Trays received water containing NaCl at three different concentrations (0, 325, and 650 mg Cl^- L^{-1}) (n = 5 of each concentration per plant species). The randomized block pattern was organized in a 91 cm x 122 cm area underneath greenhouse grow lights for a 16-hour photoperiod. Every three days, tray water was completely replaced. Chloride treatment continued for ~8 weeks at which point above ground biomass was destructively harvested by clipping where the shoot met the soil surface. Upon harvest, samples were placed in brown paper bags and dried at 65°C for 24 hours and then weighed to determine dry biomass. A subsample of each homogenized plant tissue sample was measured for chloride content at Endyne, Inc. in Williston, VT, USA[45]. Dry biomass and plant tissue Cl^- were compared across treatments for each plant species using ANOVA and post-hoc Tukey HSD (or suitable nonparametric alternatives). Duplicate soil samples per species-

treatment combination were analysed for Cl^- and electrical conductivity (EC) using the same methods as in the field study.

3. Results

3.1. Field monitoring of stormwater flows and water quality

Inlet and outlet volumes were measured during 12 storms and 20 storms over the two-year monitoring period at the Fairview and Kennedy sites, respectively. Monitored rainfall events ranged in precipitation from 0.46 to 6.20 cm for the two sites. Measured inflows into the two gravel wetlands ranged from 35 to 1627 m^3 per storm event. Inflows were strongly linearly correlated to rainfall at both Fairview ($r^2 = 0.82$, $p < 0.001$) and Kennedy ($r^2 = 0.96$, $p < 0.001$) (Figure 3). Measured outflow volumes from the two gravel wetlands ranged from 0 to 1057 m^3 per storm event and were also well correlated with rainfall at both Fairview ($r^2 = 0.66$, $p = 0.0013$) and Kennedy ($r^2 = 0.96$, $p < 0.001$) (Figure 3). At Fairview, mean event percent reductions in water volume for outflows relative to inflows were 99%, 62%, and 46% for events with rainfall < 1 cm (n = 3), 1 to 3 cm (n = 5), and > 3 cm (n = 4), respectively. Smaller volume reductions for outflows relative to inflows were observed at Kennedy, where mean reductions of 5%, 23%, and 8% occurred for events with rainfall < 1 cm (n = 7), 1 to 3 cm (n = 8), and > 3 cm (n = 5), respectively. Hydrographs for each storm monitored are shown in the Supplementary Materials. Peak flows in outflows of the SGWs were reduced on average by 48% and 83% compared to cumulative inflows at Fairview and Kennedy, respectively, across all storms monitored.

Net reduction of chloride load between inflow and outflow was frequently observed at the Fairview SGW, especially in the second year of monitoring (Figure 4a). Overall mean chloride concentrations (± 1 standard deviation) in influent and effluent observed at Fairview were in the 100-200 mg L^{-1} range both years, with no evidence of concentration reduction occurring in the SGW (Table 1). This indicates that chloride load reductions were instead linked to volume reductions at Fairview. The Kennedy SGW did not consistently act as a chloride source or sink in either year, with median load reductions closer to zero (Figure 4b) and mean influent and effluent concentrations between ~300-600 mg L^{-1} (Table 1).

Substantial reduction in TSS load between inflow and outflow, >90% removal for half of the storms monitored and >60% for over two thirds of the storms monitored, was observed at both SGWs across both years (Figure 4). For all site-year combinations, clear decreases in TSS concentrations were observed from influent to effluent (Table 1).

The SGWs acted as net P sources for approximately half of the total storms monitored across a wide range of influent P loads at both sites, and especially at the Kennedy site (Figure 5). However, the SGWs did serve as net P sinks during the five largest events monitored (inflow total P load >0.1 kg P) (Figure 5). For individual storms monitored at Fairview, median TP load reduction was positive in 2020 and negative in 2021, indicating a shift to net P export (Figure 4a). Total cumulative load reductions for TP during monitored storms were +53% and -6% in years 1 and 2, respectively, at Fairview (positive = SGW was net TP sink, negative = SGW was net TP source). At Kennedy, negative median TP load reductions were observed in both 2020 and 2021 (Figure 4b). Total cumulative TP load reductions at Kennedy for the storms monitored were -41% and +9% in years 1 and 2, respectively. The positive cumulative TP load reduction at Kennedy in year 2 was driven by three events with relatively large TP inflow (>0.1 kg P) and removal (Figure 5) – all other events that year had negative net P retention. Across all storm events monitored at both sites in both years (n = 32), overall percent reduction in TP load was +8%.

TDP results indicated that net export of dissolved P forms was a driver of net TP export, with positive mean TDP retention observed only in year 1 at Fairview (Figure 4). TDP export largely counteracted particulate P (PP) retention - across all storm events monitored at both sites in both years for both TDP and TP (n = 31), overall percent reductions in TDP and PP loads were -41% and +43%, respectively. However, our results also illustrate that reductions in PP concentration were not always observed (e.g., greater mean PP concentration in effluent compared to influent at Fairview in 2021 and at Kennedy in 2020), suggesting some net export of PP at times (Table 1). For all site-year combinations, overall mean TDP concentration was greater in the effluent than in the influent, while overall mean TP concentrations were greater in effluent than influent for three out of four site-year combinations (the exception being Fairview in 2020) (Table 1).

3.2. Field measurements of plant biomass, plant chloride, and soil chloride

Aboveground plant biomass at the two field sites ranged from 0.22-0.92 dry kg m^{-2} (Figure 6). Mean aboveground biomass (±1 standard deviation) was 0.35±0.11 dry kg m^{-2} in 2020 and 0.68±0.25 dry kg m^{-2} in 2021 at the Fairview site, and 0.28±0.07 and 0.45±0.11 dry kg m^{-2} in 2020 and 2021, respectively, at the Kennedy site. At Fairview, aboveground biomass was negatively correlated with plant tissue Cl^- ($r = -0.89$, $p = 0.018$), however no such correlation was observed at Kennedy ($p = 0.562$) (Figure 6). Plant tissue Cl^- ranged from ~8-20 mg Cl^- g^{-1} dry across both sites and years. Soil chloride did not differ significantly between years at both sites, averaging between 160 and 333 mg Cl^- kg^{-1} dry across site-year combinations (Table 2). Soil EC, however, did increase markedly in year 2 of the study relative to year 1 at both sites (Table 2).

3.3. Laboratory soil and gravel column studies

Table 3 shows general material properties for the six potential gravel wetland soil materials tested. Total solids, volatile solids, bulk density, pH, and EC ranged from 66-80 % (as-is basis), 6.7-17.7% (dry basis), 908-1015 kg m^{-3}, 5.3-7.9, and 0.1-6.5 mmhos cm^{-1}, respectively. The two native soils had lower pH and EC compared to the engineered soils. Total C ranged from 2.6-8.7% (dry basis) and total N was 0.22-0.65% (dry basis) across the six materials. Mean measured Ksat ranged widely from 0.41 to 82.70 cm hr^{-1}.

Phosphorus content and saturation varied widely across the six soil materials. Total P spanned an order of magnitude (0.03-0.37% on a dry basis), while WEP and MM-P results clearly separated the soils into relatively low P (em3, ns1, ns2) and high P (em1, em1_f, em2) groups (Table 4). The low P group had PSR_{M3} and PSR_{OX} of 0.01-0.10 and 0.03-0.09, respectively, while in the high P group PSR_{M3} was 0.75-1.34 and PSR_{OX} was 0.34-0.63 (Table 4).

For soil column experiment 1, effluent SRP concentrations for two of the engineered materials (em1 and em2) were significantly greater than both the control and the third engineered material (Dunn-Bonferroni, $p < 0.001$; Figure 7a). Mean (± 1 standard deviation) effluent SRP concentrations for the control, em1, em2, and em3 were 0.19±0.03, 0.46±0.15, 0.33±0.09, and 0.18±0.03 mg P L^{-1}, respectively, in soil column experiment 1, representing percent changes relative to the influent concentration (0.2 mg P L^{-1}) of -5%, +130%, +65%, and -10%, respectively (positive = increase in SRP concentration, negative = decrease in SRP concentration). In the

model log_{10}(SRP) ~ soil × simulation, the interaction between em1 and simulation event number was significant (p = 0.0003), indicating a trend of declining effluent SRP with successive simulations. In soil column experiment 2, no treatments had significantly different effluent SRP concentration compared to the control (Dunn-Bonferroni, p > 0.05). However, the engineered material tested after ~1 year of time in the field (em1_f, 0.23±0.06 mg P L^{-1}) did have significantly greater effluent SRP concentrations compared to the two native soils (ns1 = 0.18±0.04 mg P L^{-1}, ns2 = 0.18±0.03 mg P L^{-1}; p = 0.003 to 0.01) (Figure 7a).

Effluent TP concentration in soil column experiment 1, like SRP, was significantly greater for em1 and em2 compared to the control (Dunn-Bonferroni, p = 2.7 x 10^{-5} and 4.4 x 10^{-3}, respectively), while em3 did not differ from the control nor em1 and em2 (p > 0.05) (Figure 7b). Mean (± 1 standard deviation) effluent TP concentrations for the control, em1, em2, and em3 were 0.23±0.01, 2.37±3.07, 0.62±0.26, and 0.51±0.35 mg P L^{-1}, respectively, in soil column experiment 1, representing percent changes relative to the influent concentration (0.2 mg P L^{-1}) of +15%, +1085%, +210%, and +155%, respectively. For soil column experiment 2, TP effluent for the engineered soil collected after 1 year in the field (em_f, 0.42±0.22 mg P L^{-1}) was significantly greater than that for the control and both native soils (Dunn-Bonferroni, p < 0.025 in all cases), which were all within 12.5% of the influent 0.2 mg P L^{-1} on average (Figure 7b). Linear models of the form log_{10}(TP) ~ soil × simulation indicated significant interactions between em1 and simulation number (p = 0.0001), as well as em1_f and simulation number (p = 0.0251), indicating trends of declining effluent TP with successive simulations for these two soils.

TSS concentrations in soil column effluents for em1, em2, and em3 were all significantly greater than the control in experiment 1 (p = 0.005 to 0.050), with means ranging from 18-69 mg TSS L^{-1} compared to 2 mg L^{-1} for the control effluent (Figure 7c). In the model log_{10}(TSS+1) ~ soil × simulation, the interaction between em1 and simulation event number was significant (p = 0.042), indicating a trend of declining effluent TSS with successive simulations. In experiment 2, TSS was variable for all treatments with no significant differences from the control detected (p > 0.08 in all cases) (Figure 7c).

In the gravel column experiment, influent water was characterized by overall mean values of 0.25 mg SRP L^{-1}, 0.56 mg TP L^{-1}, and 21 mg TSS L^{-1}. Significantly reduced effluent concentrations of SRP were observed for granite and limestone in comparison to the control (Dunn-Bonferroni, p = 0.014 for granite, p = 0.0001 for limestone) (Table 5). However, TP effluent decrease relative to the control was only significant for limestone (Tukey HSD, p = 0.0389), and the effluent TP concentrations across all three gravels tested were 25-70% greater than the initial concentration used in the synthetic stormwater at the beginning of the linked soil column experiment (0.20 mg P L^{-1}) (Table 5). TSS concentrations in gravel column effluents were variable (mean = 4-15 mg TSS L^{-1}) and did not differ across all treatments, including the control (Table 5). For all gravel columns, including the control, pH was greater in the effluent (medians = 7.4, 7.8, 8.0, and 8.0 for control, granite, quartzite, and limestone, respectively) than influent (medians = 7.1-7.2).

For all lab column experiments, chloride concentrations in effluents from soils and gravels (632-765 mg Cl^- L^{-1}) remained similar to influents (608-769 mg Cl^- L^{-1}), and near the initial synthetic stormwater concentration of ~650 mg Cl^- L^{-1}, with no indication of Cl^- removal or net export for any column.

3.3. Laboratory chloride bioassay

Broad-fruited bur-reed biomass after 8 weeks did not significantly differ across chloride treatments (ANOVA, p = 0.112) (Table 6). Shallow sedge biomass, however, declined with increased chloride, with biomass for the 650 mg Cl^- L^{-1} treatment significantly lower than that in the 0 mg Cl^- L^{-1} treatment (Dunn-Bonferroni, p = 0.022) (Table 6). Plant tissue chloride at the conclusion of the 8-week bioassay indicated effects of the chloride treatment, with increasing chloride in plant tissues as the chloride concentration in the irrigation water increased (Table 6). Soil chloride content at the end of the bioassay indicated accumulation of Cl^- with greater Cl^- concentration in the irrigation water for both broad-fruited bur-reed (mean = 293, 5173, and 9025 mg Cl^- kg^{-1} for the 0, 325, and 650 mg Cl^- L^{-1} treatments, respectively) and shallow sedge (mean = 139, 3508, and 5675 mg Cl^- kg^{-1} for the 0, 325, and 650 mg Cl^- L^{-1} treatments, respectively). Soil EC also increased for treatments with Cl^- added to the irrigation water (0.73-1.63 mmhos cm^{-1}) relative to the 0 mg Cl^- L^{-1} treatment (0.18-0.29 mmhos cm^{-1}).

4. Discussion

4.1. SGW flow attenuation and P retention in the field

Both SGWs included in our field study are located in flow-impaired watersheds. Therefore, these systems were required to attenuate flow from the Channel Protection Storm (1-yr, 24-hr rainfall depth)[3] which produces a total rainfall of 5.3 cm. Flow monitoring of the two SGWs in the first two years of their construction provided empirical evidence that these systems were effective at attenuating peak flows, with average percent reductions in peak flow of 48% at Fairview and 83% at Kennedy across storms monitored in both sampling years (Table S1, Supplementary Materials). For storms similar in size to the Channel Protection Storm, peak flow reduction was 11% at Fairview and 81% at Kennedy across both years, indicating that Kennedy better reduced flow rates for such events. We also commonly observed flow volume reductions at both sites across a range of inflow amounts, and especially at Fairview (Figure 3). This result was surprising, as these SGWs were designed to attenuate peak flows, but not to reduce overall flow volumes, due to a permanently saturated gravel layer. The observed flow volume reductions indicate that infiltration may be occurring at times, freeing up storage space in the gravel layer between storm pulses.

Our field results demonstrate that, despite net P retention for several individual storm events, both SGWs failed to achieve the desired P load reductions of 60-80% (Figures 4 and 5). Instead, overall load reductions in TP across all storms monitored in both systems during this two-year study revealed minor P retention (+8% load reduction), with differences in performance between systems and years (Figure 4). These results are more similar to those found by Sullivan and McDonald[10] than the much more promising results from the University of New Hampshire Stormwater Center[7-9]. Furthermore, our field monitoring results illustrated a trade-off between particulate P capture and dissolved P export, with each mechanism being of approximately the same magnitude overall with opposite effects on net TP load retention (Figure 4), although concentration data indicated that particulate P concentration was sometimes greater in outflows compared to inflows (Table 1). Therefore, these SGWs can be effective at particle trapping, but susceptible to dissolved P loss. Several interrelated factors likely contributed to the observed net dissolved P export, including insufficient P sorption capacity of materials used in the SGWs, P leaching from the soil layer, and dissolved P released upon the mineralization of plant litter. The

latter mechanism may help explain why we observed positive mean net retention of TDP in Year 1 at Fairview when vegetation was relatively quick to establish (and likely assimilated some of the P provided by the soil media), but then net TDP export in Year 2 after Year 1 plants had senesced over winter. Vegetation in the SGW system is ultimately limited in its capacity to assimilate P into biomass and a portion of the P stored in biomass one year can be released in subsequent years via mineralization[46].

4.2. Effects of SGW material selection on performance and recommendations for future material selection

Collectively, the soil column testing results (Figure 7) indicate that two out of the three engineered materials acquired from compost/soil manufacturers and tested here will likely be a source of SRP loss post-installation in the field (em1 and em2). Given that em1 was the material used at our two field sites, the lab results suggest that P leaching from the soils installed at the two SGWs studied in the field was very likely a key factor leading to the net export of TDP that we observed. SRP loss from engineered soils will likely decline over time, as evidenced by the SRP results for em1 (reduced SRP in effluent across successive storms) and em1_f (not different from control), which originated from the same source, but with the latter analysed approximately one year after field installation. SRP leaching was not observed for native soils collected from two SGW field sites (ns1 and ns2), indicating that utilizing native soils has advantages over engineered materials from a P perspective (as well as from a cost perspective). Fine solids with attached P are also likely to be lost from all three engineered materials tested here, adding to the overall P loss from the material on top of SRP leaching, although this loss of fine particles could fade over time, as evidenced by the TSS results for em1_f (not different from control).

P leaching from compost-based materials in green stormwater infrastructure has been documented by previous researchers as well. Mangum et al.[47] reported that compost-amended gravel wetland mesocosms leached P above that of mesocosms using a soil media without compost, leading to maximum P concentrations of 2.9 mg P L^{-1} for 30% compost and 0.52 mg P L^{-1} for 15% compost. Hurley et al.[21] investigated the effect of saturation duration on SRP leaching from composts and compost-amended bioretention mixes, finding increasing P leaching with increasing saturation time and greater SRP concentrations in compost leachate compared to bioretention mixes. Clear design specifications are needed to help guide effective use of compost and soil materials in green infrastructure (including both SGWs and bioretention) that mitigates unintended P leaching.

Based on the results of our laboratory soil column tests (Figure 7), soil characterization results (Table 4), and existing literature, we recommend that any soils (or final soil mixes) used in SGWs and other green stormwater infrastructure should have a Mehlich-3 PSR (PSR_{M3}) no greater than 0.10. Several soil studies have reported thresholds near 0.10 for PSR_{M3}, above which release of SRP is more likely to occur[34,35]. For SGW soils assessed in our column leaching tests, those with $PSR_{M3} \leq 0.10$ did not leach any SRP (Figure 7, Table 4). PSR_{OX} based on oxalate-extractable P, Al, and Fe is also an effective metric for gauging soil P release risk, however it is also less routinely offered by commercial soils labs compared to Mehlich-3 extraction. Wiegman et al.[48] reported a threshold in PSR_{OX} of 0.23 for riparian soils in Vermont, above which SRP release increases. In our study, all potential SGW soils having $PSR_{M3} \leq 0.10$ also had $PSR_{OX} < 0.10$ (Table 4) and therefore are well below the PSR_{OX} threshold defined by Wiegman et al.[48]. Based on this combination of evidence, our recommended limit for PSR_{M3} should be effective in greatly reducing the potential for dissolved P export from SGWs and other green stormwater infrastructure in the field upon installation.

One caveat for the recommended PSR_{M3} limit described above is that it is possible that a soil material testing at or below $PSR_{M3} = 0.10$ could be a source of other non-SRP forms of P. For example, in the lab study, engineered material #3 (em3) was a net sink for SRP, but a net source of total P (Figure 7), indicating that other P forms, such as particulate inorganic P, particulate organic P, or dissolved organic P, were exported from the em3 material. Such non-SRP P export will likely be most pronounced immediately following installation as fine particles are transported out of the soil layer.

In Vermont, designers including some authors of this paper have started to test materials and native soils, applying our recommended PSR_{M3} limit of 0.10. Finding soil materials that meet this specification has been manageable in these cases. This includes some native soils at SGW sites, as well as mixtures of local clay and peat. We expect that meeting desired saturated hydraulic conductivity (K_{sat}) will be more challenging when sourcing soils for SGWs. Previous recommendations used by practitioners in Vermont and elsewhere in the U.S. for SGWs have specified that the soil material shall have K_{sat} in the range of 0.013 to 0.13 cm hr^{-1} (0.01 - 0.1 ft d^{-1}). Assuming low to medium bulk density, these values are near the range of a soil with a textural analysis of 'clay', 'silty clay', 'clay loam', or 'silty clay loam', and possibly 'sandy clay' and 'silty loam' on the high end[49]. None of the engineering materials or native soils tested here (K_{sat} = 0.41 to 82.70 cm hr^{-1}) met this goal (Table 3). This means SGWs employing any of the soil materials tested in this study will likely have more infiltration through the soil layer than desired, and sometimes much more, which could lead to short-circuiting in the system. More research is needed to guide use of soils that both limit potential for P leaching and meet hydraulic goals, including monitoring to determine the potential for fines from clayey soils to migrate into the gravel layer and cause clogging. We also recommend that SGW designs be studied that forego use of a soil layer altogether. However, such designs will need to ensure that horizontal subsurface flow through the gravel layer is maintained.

Limited gravel P sorption capacity is another potential driver of the poor TDP retention that we observed in our field study. Previous studies have reported highly variable P sorption capacity for different gravel materials, which are characterized by different surface areas and chemical properties[50-53]. The results of our lab study indicated that the gravels tested provide some P sorption capacity, but this is not enough to completely counteract the loss of P from the engineered soil. Mean TP effluent concentrations for granite, quartzite, and limestone columns were 0.30, 0.34, and 0.25 mg P L^{-1}, equal to 150%, 170%, and 125% of the P concentration for the synthetic stormwater solution initially fed to the soil columns "upstream" in the lab tests (0.20 mg P L^{-1}). This illustrates the potential for SGW substrates (soil + gravel) to cumulatively serve as net P sources, despite evidence of P retention by granite and limestone (according to SRP results) and especially limestone (according to SRP and TP results) (Table 5).

Ultimately, decreasing direct P leaching risk from the soil layer upon installation may not be enough to ensure long-term mitigation of SGW dissolved P loss due to ongoing P loading, inherent limitations in P assimilation by vegetation, P release from vegetation - which may be exacerbated by freeze-thaw cycles in cold regions[54], and limited SGW P sorption capacity. There may also be opportunities to use geochemical augmentation to increase the P sorption capacity of the soil layer and thus the SGW. For example, aluminium-based

drinking water treatment residuals have been shown to enhance the P sorption capacity of bioretention media[55]. Therefore, these materials or others with high P sorption capacity should be investigated as an ingredient in SGW soil mixtures to reduce PSR_{M3}, increase capacity to sorb dissolved P loaded directly or released from plant litter via mineralization, and in turn reduce the risk of dissolved P export long-term.

4.3. Chloride movement and storage in SGWs in the field

Our field study results showed substantial loading of chloride to the two urban SGWs, especially at the Kennedy site where influent Cl^- concentrations were 409 ± 471 mg L^{-1} in year 1 and 306 ± 257 in year 2 (Table 1). Due to the highly impervious and commercialized land use within the Kennedy Drive wetland's drainage area, salt applications are a common practice in the winter months as a road de-icing mechanism. There was evidence of positive net chloride retention at Fairview, but not at Kennedy (Figure 4). This difference was likely due to three factors at Fairview – lesser influent chloride loads, greater flow volume reductions (Figure 3), and better vegetation establishment during the study period (Figures S1 and S2, Supplemental Materials). Wetland vegetation provided a chloride storage mechanism in the SGWs, but much like plant P storage, this is ultimately finite and chloride assimilated into vegetation one year may be susceptible to release in future years. The observed inverse relationship between plant biomass and plant tissue chloride at Fairview (Figure 6) provides some limited evidence of chloride suppression of vegetation growth (which likely reduces plant assimilation of P). Plant tissue chloride at Kennedy was more likely to be ≥14 mg g^{-1}, and the greater chloride concentration of influent stormwater (Table 1) was possibly one factor in the slower establishment of vegetation at that site. Analysis of SGW soils in both years did not reveal obvious storage of chloride in the soil layer, although the observed increase in soil EC between years at both sites did indicate soil salinization is occurring over time (Table 2). Aside from negative impacts on vegetation[25,26], this could also decrease the soil permeability[23], which may be advantageous in SGWs where infiltration through the soil layer is not desirable, unlike in the context of bioretention. Some researchers have hypothesized that the increased ionic strength associated with salinization should weaken P sorption in freshwater soils and sediments (e.g., via Cl^- competing with PO_4^{3-} for sorption sites), leading to P loss[56]. However, experiments testing this hypothesis have shown conflicting results, potentially due to concomitant changes with salinization that could instead enhance P sorption depending on the degree of salinization, the ionic composition of salinizing water, and soil/sediment characteristics[56].

4.4. Effects of SGW material selection on chloride transport and potential effects of chloride on SGW vegetation

The laboratory column experiment results, which showed no retention of chloride in soils or gravels, confirmed that SGW media should be expected to have negligible direct impacts on chloride transport in the field. The laboratory bioassay results illustrated there may be an indirect, minor, and likely temporary chloride storage mechanism – evaporation of water from soils resulted in soil chloride accumulation. We also observed different responses to chloride loading in the bioassay by two plant species commonly used in wetland seed mixes in the Northeast U.S. Shallow sedge (*Carex lurida*) growth was inhibited by increasing chloride in the irrigation water, whereas this was not the case for Broad-fruited bur-reed (*Sparganium eurycarpum*) (Table 6). Variation of tolerance among species may provide a competitive advantage in the field to species more tolerant of chloride, including invasive species like reed canary grass (*Phalaris arundinacea*)[57,58]. Using high salt accumulating plants to treat chloride is a form of phytoremediation known as phytodesalinization[59]. This is a process that has been studied under a number of varying conditions and with different species, however there is limited information on the comparative potential of different species in the context of constructed wetlands[60-65]. Our lab bioassay results provide additional evidence alongside the field results that assimilation of chloride by vegetation provides a chloride storage mechanism (Table 6). However, even at relatively high foliar chloride content (e.g., 35 mg Cl^- g^{-1} dry tissue) and plant dry biomass (e.g., 1 kg m^{-2}), this potential storage of chloride in SGW aboveground biomass (e.g., ~35 g Cl^- m^{-2}) can be much smaller than the incoming load of chloride (e.g., 2306 kg Cl^- for the 20 storms monitored at the Kennedy SGW, or ~6800 g Cl^- m^{-2}).

5. Conclusions

Our field study indicated minor overall net TP retention by SGWs treating urban stormwater, falling far short of expected performance. Instead, the SGWs were characterized by a trade-off between particulate P capture and dissolved P export, with each being of a similar magnitude. Our laboratory experiment demonstrated that the dissolved P export observed in the field is likely driven by P leaching from the engineered soil material used (em1), although limited P sorption capacity and mineralization of P temporarily stored in vegetation are also likely factors. Gravel materials in SGWs are unlikely to provide sufficient P sorption capacity to offset P leaching from engineered soil materials with relatively high P saturation. Therefore, we recommend that any engineered soil materials or native soils used in SGWs be characterized by $PSR_{M3} \leq 0.10$ to greatly reduce P leaching risk. This recommendation can also apply to other types of green stormwater infrastructure designed for phosphorus retention. It is critical that soils selected for SGWs also have sufficiently low hydraulic conductivity, otherwise undesirable infiltration through the soil layer can occur resulting in short circuiting. In cold climates, high chloride loads driven by road salting will likely affect SGWs over time. Our results collectively indicate that chloride loads will provide competitive advantage to some plant species over others, that plant assimilation can serve as Cl^- sink – albeit one that is temporary and small compared to incoming Cl^- loads over time, and that SGW media (soils and gravels) likely have minimal effects on chloride transport. More experimentation and monitoring is needed to inform best practices for SGW design to effectively treat urban stormwater in cold climates. Our results suggest that the increasing popularity of SGWs in Vermont and other Northeast USA locations will likely make only minimal contributions to watershed P load reduction goals, and perhaps even have adverse effects, unless designs are improved to enhance P removal and be resilient to chloride loading. Our findings can also inform SGW design and use of engineered soils in urban stormwater BMPs more broadly.

Author Contributions

Conceptualization (E.D.R., A.T., D.R.), Data Curation (E.D.R., N.N., M.K.), Formal Analysis (E.D.R., A.T., N.N., M.K.), Funding Acquisition (E.D.R., A.T., D.R.), Investigation (E.D.R., N.N., M.K., T.A., M.Y., F.A.B., A.R.H.W.), Methodology (E.D.R., A.T., N.N., M.K., F.A.B.,

A.R.H.W., D.R.), Project Administration (E.D.R., A.T.), Resources (E.D.R., A.T.), Supervision (E.D.R., A.T., D.R.), Validation (E.D.R., A.T.), Visualization (E.D.R., N.N., M.K.), Writing – Original Draft Preparation (E.D.R., A.T., N.N., M.K., T.A., M.Y.), Writing – Review & Editing (E.D.R., A.T., N.N., M.K., T.A., M.Y., F.A.B., A.R.H.W., D.R.).

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work received support from the National Oceanic and Atmospheric Administration (NOAA) National Sea Grant College Program Award NA18OAR4170099 to the Lake Champlain Sea Grant Institute. We acknowledge the following individuals for their help with data collection, analysis, equipment repair, and field monitoring configuration: Chelsea Mandigo, Dave Wheeler, Giuliana Frizzi, James Eyler, Jessica Rubin, Michaela Francesconi, Noah McAllister, Ryan Leo, and Steve McManus.

References

[1] Center for Watershed Protection, 2007, National Pollutant Removal Performance Database, Version 3. Ellicott City, MD.

[2] UNH Stormwater Center, 2022, UNHSC Subsurface Gravel Wetland Design Specifications. https://scholars.unh.edu/cgi/viewcontent.cgi?article=1073&context=stormwater

[3] Vermont Agency of Natural Resources, 2017, 2017 Vermont stormwater management manual rule and design guidance. https://dec.vermont.gov/sites/dec/files/wsm/stormwater/docs/Permitinformation/2017%20VSMM_Rule_and_Design_Guidance_04172017.pdf

[4] US EPA, 2016, Phosphorus TMDLs for Vermont segments of Lake Champlain.

[5] J. Garcia, D.P.L. Rousseau, J. Morato, E.L.S. Lesage, V. Matamoros, and J.M. Bayona, Contaminant removal processes in subsurface-flow constructed wetlands: a review, *Crit. Rev. Environ. Sci. Technol.*, 2010, **40**, 561–661.

[6] J. Vymazal, Constructed wetlands for wastewater treatment: five decades of experience, *Environmental Science & Technology*, 2011, **45**, 61-69.

[7] J.J. Houle and T.P. Ballestero, 2020, Some Performance Characteristics of Subsurface Gravel Wetlands for Stormwater Management. In World Environmental and Water Resources Congress 2020: Water, Wastewater, and Stormwater and Water Desalination and Reuse (pp. 125-135). Reston, VA: American Society of Civil Engineers.

[8] T.P. Ballestero, J.J. Houle, and T.A. Puls, 2016, Breaking Through: University Of New Hampshire Stormwater Center Report.

[9] R.M. Roseen, T.P. Ballestero, J.J. Houle, P. Avellaneda, J. Briggs, G. Fowler, and R. Wildey, Seasonal performance variations for storm-water management systems in cold climate conditions, *Journal of Environmental Engineering*, 2009, **135**, 128-137.

[10] C. Sullivan and W. McDonald, Hydrologic and water quality performance of a subsurface gravel wetland treating stormwater runoff, *Journal of Environmental Management*, 2022, **322**, 116120.

[11] The Water Research Foundation, 2020, International Stormwater BMP Database: 2020 Summary Statistics. https://www.waterrf.org/system/files/resource/2020-11/DRPT-4968_0.pdf

[12] US EPA, 2009, Stormwater wet pond and wetland management guidebook. EPA 833-B-09-001.

[13] T.R. Schueler, 2000, The environmental impact of stormwater ponds. Center for Watershed Protection. 443–452.

[14] A.S. Chiandet and M.A. Xenopoulos, Landscape controls on seston stoichiometry in urban stormwater management, *Freshwater Biology*, 2011, **56**, 519–529.

[15] S. Duan, T. Newcomer-Johnson, P. Mayer, and S. Kaushal, Phosphorus retention in stormwater control structures across streamflow in urban and suburban watersheds, *Water*, 2016, **8**, 390.

[16] N.A. McEnroe, J.M. Buttle, J. Marsalek, F.R. Pick, M.A. Xenopoulos, and P.C. Frost, Thermal and chemical stratification of urban ponds: Are they 'completely mixed reactors'?, *Urban Ecosystems*, 2013, 16, 327–339.

[17] K. Song, M. Xenopoulos, J. Marsalek, and P.C. Frost, The fingerprints of urban nutrients: dynamics of phosphorus speciation in water flowing through developed landscapes, *Biogeochemistry*, 2015, **125**, 1–10.

[18] K. Song, M.A. Xenopoulos, J. Buttle, J. Marsalek, N. Wagner, F. Pick, and P.C. Frost, Thermal stratification patterns in urban ponds and their relationships with vertical nutrient gradients, *Journal of Environmental Management*, 2013, **127**, 317-323.

[19] G.H. Le Fevre, K.H. Paus, P. Natarajan, J.S. Gulliver, P. Novak, and R.M. Hozalski, Review of dissolved pollutants in urban storm water and their removal and fate in bioretention cells, *J. Environ. Eng.*, 2015, **141**, 04014050.

[20] J. Li and A.P. Davis, A unified look at phosphorus treatment using bioretention, *Water Res.*, 2016, **90**, 141–155.

[21] S. Hurley, P. Shrestha, and A. Cording, Nutrient leaching from compost: Implications for bioretention and other green stormwater infrastructure, *Journal of Sustainable Water in the Built Environment*, 2017, **3**, 04017006.

[22] P. Shrestha, S.E. Hurley, and B.C. Wemple, Effects of different soil media, vegetation, and hydrologic treatments on nutrient and sediment removal in roadside bioretention systems, *Ecological Engineering*, 2018, **112**, 116-131.

[23] S.P. Kakuturu and S.E. Clark, Clogging mechanism of stormwater filter media by NaCl as a deicing salt, *Environmental Engineering Science*, 2015, **32**, 141-152.

[24] W.D. Hintz and R.A. Relyea, A review of the species, community, and ecosystem impacts of road salt salinization in fresh waters, *Freshwater Biology*, 2019, **64**, 1081-1097.

[25] J.A. Richburg, W.A. Patterson, and F. Lowenstein, Effects of road salt and *Phragmites australis* invasion on the vegetation of a western Massachusetts calcareous lake-basin fen, *Wetlands*, 2001, **21**, 247–255.

[26] P.J. White and M.R. Broadley, Chloride in soils and its uptake and movement within the plant: a review, *Annals of Botany*, 2001, **88**, 967–988.

[27] S.S. Kaushal, P.M. Groffman, G.E. Likens, K.T. Belt, W.P. Stack, V.R. Kelly, L.E. Band, and G.T. Fisher, Increased salinization of fresh water in the northeastern United States, *Proceedings of the National Academy of Sciences*, 2005, **102**, 13517-13520.

[28] C.R. Burgis, G.M. Hayes, D.A. Henderson, W. Zhang, and J.A. Smith, Green stormwater infrastructure redirects deicing salt from surface water to groundwater, *Science of the Total Environment*, 2020, **729**, 138736

[29] D. Skultety and J.W. Matthews, Urbanization and roads drive non-native plant invasion in the Chicago Metropolitan region, *Biol. Invasions*, 2017, **19**, 2553–2566.

[30] Y.C. Li, D.Q. Zhang, and M. Wang, Performance evaluation of a full-scale constructed wetland for treating stormwater runoff, *CLEAN – Soil, Air, Water*, 2017, **45**, 160074.

[31] N.E. Scott, A. W. Davison, De-icing salt and the invasion of road verges by maritime plants, *Watsonia*, 1982, **14**, 41–52.

[32] B. Hoskins, 1997, Soil testing handbook for professionals in agriculture, horticulture, nutrient and residuals management. University of Maine. https://umaine.edu/soiltestinglab/wp-content/uploads/sites/227/2016/07/handbook.pdf

[33] F. Courchesne and M.C. Turmel. 2007. Extractable Al, Fe, Mn, and Si. In: M.R. Carter and E.G. Gregorich, Soil sampling and methods of analysis. 2nd ed. CRC Press, Boca Raton, FL. p. 307–315.

[34] V.D. Nair, Soil phosphorus saturation ratio for risk assessment in land use systems, *Frontiers in Environmental Science*, 2014, **2**, 6.

[35] B. Dari, V.D. Nair, A.N. Sharpley, P. Kleinman, D. Franklin, and W.G. Harris, Consistency of the threshold phosphorus saturation ratio across a wide geographic range of acid soils, *Agrosystems, Geosciences & Environment*, 2018, **1**, 1-8.

[36] P. Kleinman, D. Sullivan, A. Wolf, R. Brandt, Z. Dou, H. Elliot, J. Kovar, A. Leytem, R. Maguire, P. Moore, and L. Saporito, Selection of a water-extractable phosphorus test for manures and biosolids as an indicator of runoff loss potential, *Journal of Environmental Quality*, 2007, **36**, 1357-1367.

[37] C.J. Penn, M.R. Williams, J. Camberato, N. Wenos, and H. Wason, Desorption kinetics of legacy soil phosphorus: Implications for non-point transport and plant uptake, *Soil Systems*, 2022, **6**, 6.

[38] E. D'Angelo, J. Crutchfield, and M. Vandiviere, Rapid, sensitive, microscale determination of phosphate in water and soil, *Journal of Environment Quality*, 2001, **30**, 2206-2209.

[39] T. Ohno and L.M. Zibilske, Determination of low concentrations of phosphorus in soil extracts using malachite green, *Soil Science Society of America Journal*, 1991, **55**, 892-895.

[40] S. Rahutomo, J.L. Kovar, and M.L. Thompson, Malachite green method for determining phosphorus concentrations in diverse matrices, *Communications in Soil Science and Plant Analysis*, 2019, **50**, 1743-1752.

[41] C.J. Patton, J.R. Kryskalla, 2003, Methods of analysis by the U.S. Geological Survey national water quality laboratory—Evaluation of alkaline persulfate digestion as an alternative to kjeldahl digestion for determination of total and dissolved nitrogen and phosphorus in water. Water Resources Investigations Report 03(4174).

[42] J. Murphy and J.P. Riley, A modified single solution method for determination of phosphate in natural waters, *Analytica Chimica Acta*, 1962, **27**, 31–36.

[43] E.D. Roy, E.A. Smith, S. Bargu, and J.R. White, Will Mississippi River diversions designed for coastal restoration cause harmful algal blooms?, *Ecological Engineering*, 2016, **91**, 350–364.

[44] R.L. Powell, R.A. Kimerle, and E.M Moser, Development of a plant bioassay to assess toxicity of chemical stressors to emergent macrophytes, *Environmental Toxicology and Chemistry: An International Journal*, 1996, **15**, 1570-1576.

[45] U.S. Environmental Protection Agency, 1986, Method 6010: Inductivity coupled plasma atomic emission spectroscopy. In Test Methods for Evaluating Solid Waste. Vol. 1A: Laboratory Manual Physical/Chemical Methods. SW-846. Washington, DC.

[46] K.R. Reddy, R.D. DeLaune, 2008, Biogeochemistry of wetlands: science and applications. CRC press.

[47] K.R. Mangum, Q. Yan, T.K. Ostrom, and A.P. Davis, *Journal of Environmental Engineering*, 2020, **146**, 04019128.

[48] A.R. Wiegman, G.H. Myers, I.C. Augustin, M.L. Kubow, M.J. Fein-Cole, V.L. Perillo, D.S. Ross, R.M. Diehl, K.L. Underwood, W.B. Bowden, and E.D. Roy, *Biogeochemistry*, 2022, **161**, 137-156.

[49] USDA NRCS, 2010, Part 618 – Soil properties and qualities. In: Title 430 – National Soil Survey Handbook.

[50] C.Y. Lee, C.C. Lee, F.Y. Lee, S.K. Tseng, and C.J. Liao, Performance of subsurface flow constructed wetland taking pretreated swine effluent under heavy loads, *Bioresource Technology*, 2004, **92**, 173-179.

[51] G. Xu, Y. Li, W. Hou, S. Wang, and F. Kong, Effects of substrate type on enhancing pollutant removal performance and reducing greenhouse gas emission in vertical subsurface flow constructed wetland, *Journal of Environmental Management*, 2021, **280**, 111674.

[52] H. Lu, L. Xiao, T. Wang, S. Lu, H. Wang, X. Guo, and J. Li, The application of steel slag in a multistage pond constructed wetland to prify low-phosphorus polluted river water, *Journal of Environmental Management*, 2021, **292**, 112578.

[53] C. Vohla, M. Kõiv, H.J. Bavor, F. Chazarenc, and Ü Mander, *Ecological Engineering*, 2011, **37**, 70-89.

[54] C.J. Whitfield, N.J. Casson, R.L. North, J.J. Venkiteswaran, O. Ahmed, J. Leathers, K.J. Nugent, T. Prentice, and H.M. Baulch, The effect of freeze-thaw cycles on phosphorus release from riparian macrophytes in cold regions, *Canadian Water Resources Journal*, 2019, **44**, 160-173.

[55] M.R. Ament, S.E. Hurley, M. Voorhees, E. Perkins, Y. Yuan, J.W. Faulkner, and E.D. Roy, *ACS ES&T Water*, 2021, **1**, 688-697.

[56] L. Kinsman-Costello, E. Bean, A. Goeckner, J.W. Matthews, M. O'Driscoll, M.M. Palta, A.L. Peralta, A.J. Reisenger, G.J. Reyes, A.R. Smyth, and M. Stofan, Mud in the city: Effects of freshwater salinization on inland urban wetland nitrogen and phosphorus availability and export, *Limnology and Oceanography Letters*, **2023**, 8, 112-130.

[57] A. Shambaugh, 2008, Environmental implications of increasing chloride levels in Lake Champlain and other basin waters. Vermont Agency of Natural Resources, Department of Environmental Conservation

[58] M. Schück and M. Greger, Chloride removal capacity and salinity tolerance in wetland plants, *Journal of Environmental Management*, 2022, **308**, 114553.

[59] M.S., Fountoulakis, G. Sabathianakis, I. Kritsotakis, E.M. Kabourakis, and T. Manios, Halophytes as vertical-flow constructed wetland vegetation for domestic wastewater treatment, *Science of the Total Environment*, 2017, **583**, 432-439.

[60] A.J. Lymbery, R.G. Doupé, T. Bennett, and M.R. Starcevich, Efficacy of a subsurface-flow wetland using the estuarine sedge *Juncus kraussi* to treat effluent from inland saline aquaculture, *Aquacultural Engineering*, 2006, **34**, 1-7.

[61] B. Morteau, G. Triffault-Bouchet, R. Galvez, L. Martel, and S. Leroueil, Treatment of salted road runoffs using *Typha latifolia*, *Spergularia canadensis*, and *Atriplex patula*: a comparison of their salt removal potential, *Journal of ASTM International*, 2009, **6**, 1-7.

[62] S. Nilratnisakorn, P. Thiravetyan, and W. Nakbanpote, A constructed wetland model for synthetic reactive dye wastewater treatment by narrow-leaved cattails (*Typha angustifolia* Linn.), *Water Science and Technology*, 2009, **60**, 1565-1574.

[63] E.R. Rozema, R.J. Gordon, and Y. Zheng, Plant species for the removal of Na^+ and Cl^- from greenhouse nutrient solution, *HortScience*, 2014, **49**, 1071-1075.

[64] E.R. Rozema, R.J. Gordon, and Y. Zheng. Harvesting plants in constructed wetlands to increase biomass production and Na^+ and Cl^- removal from recycled greenhouse nutrient solution, *Water, Air, & Soil Pollution*, 2016, **227**, 136.

[65] O. Shelef, A. Gross, and S. Rachmilevitch, The use of *Bassia indica* for salt phytoremediation in constructed wetlands, *Water Research*, 2012, **46**, 3967-3976.

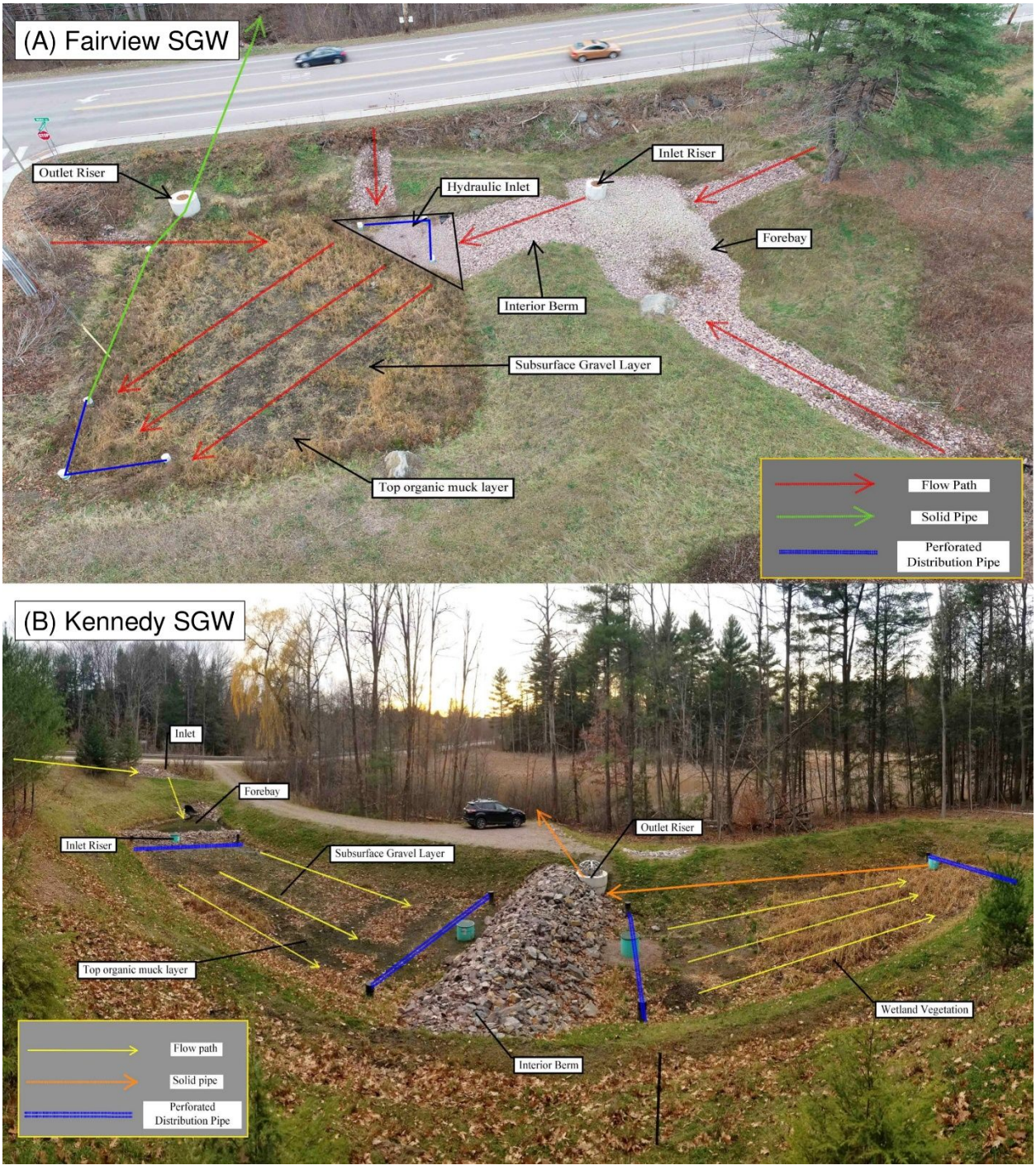

Figure 1. Diagrams of the (a) Fairview and (b) Kennedy stormwater subsurface gravel wetlands (SGW) monitored during the field study. Water primarily moves through the SGWs by movement into perforated inlet riser and distribution pipes followed by subsurface horizontal flow through a gravel layer controlled by an outlet structure to remain permanently saturated. See main text for additional site descriptions.

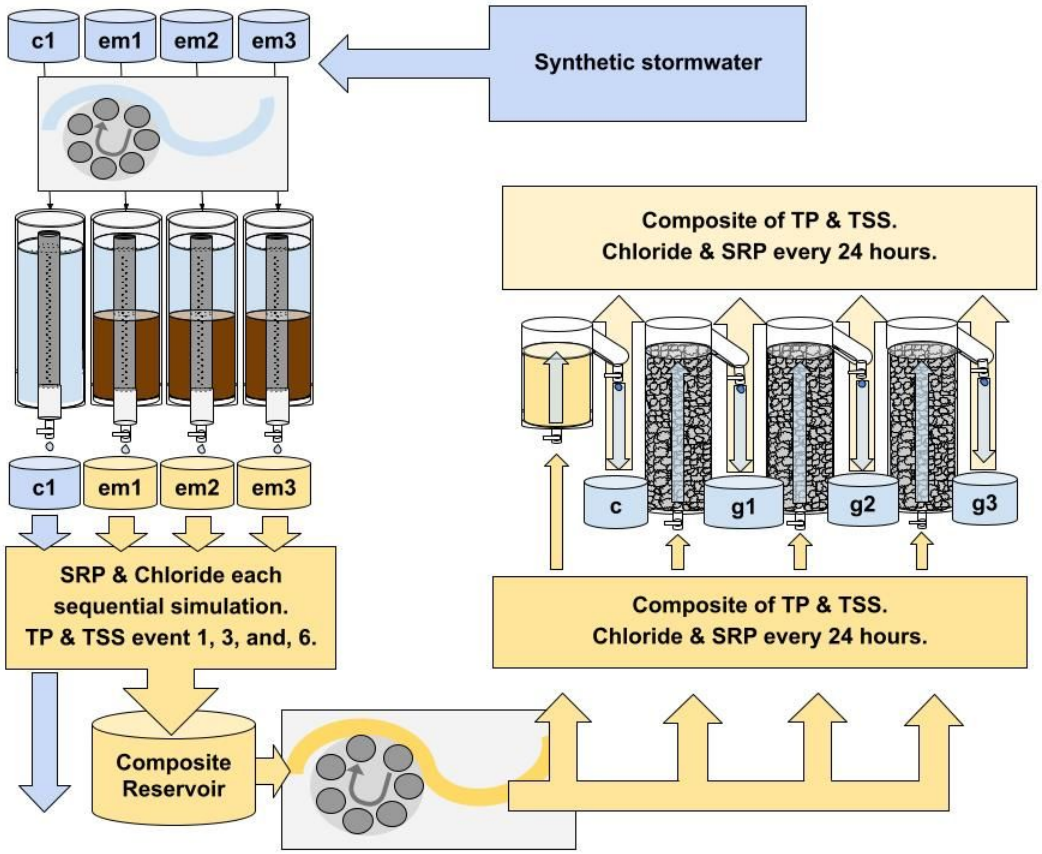

Figure 2. Flow diagram for methods used in soil columns experiment 1 and the linked gravel columns experiment. Control columns are denoted by c1 and c. Soils included three engineered materials (em1, em2, em3), and three gravels were tested (g1 = granite, g2 = quartzite, g3 = limestone). Not pictured: an additional engineered material (em1_f) and two native soils (ns1 and ns2) were tested with a second control (c2) using the same soil column methods shown on the left side of this figure.

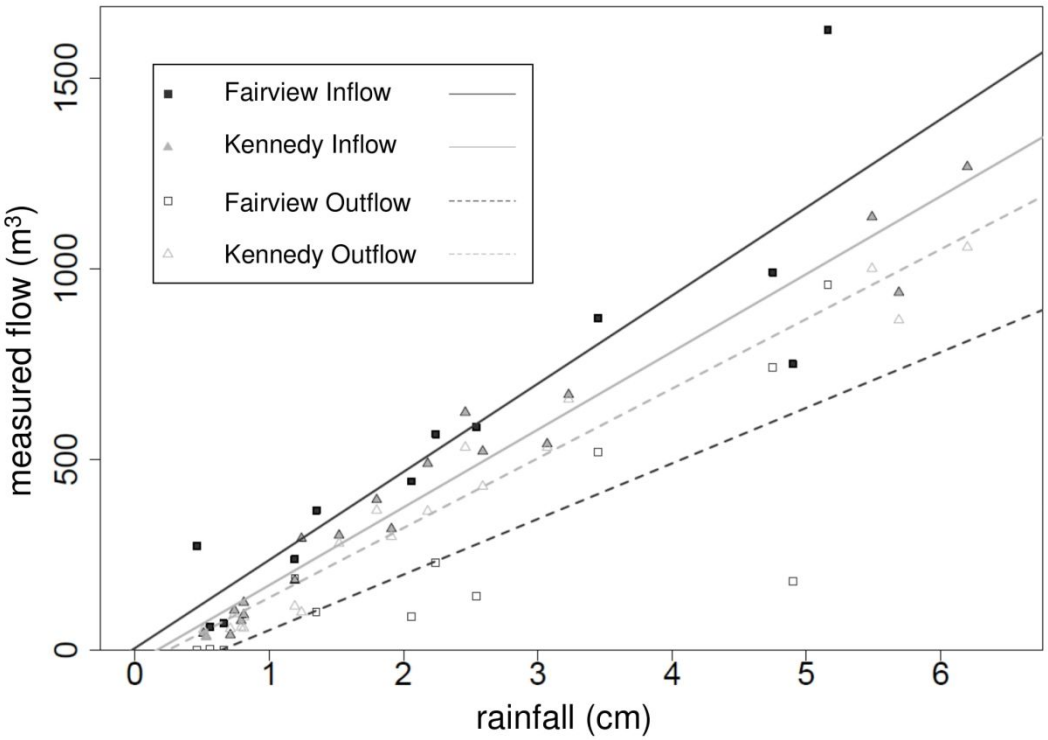

Figure 3. Measured flows (m^3) for storms monitored at the two stormwater subsurface gravel wetland field sites as a function of storm rainfall amount (cm). Solid and dashed lines represent simple linear regression fits for inflows and outflows, respectively. See text for descriptions of statistical results.

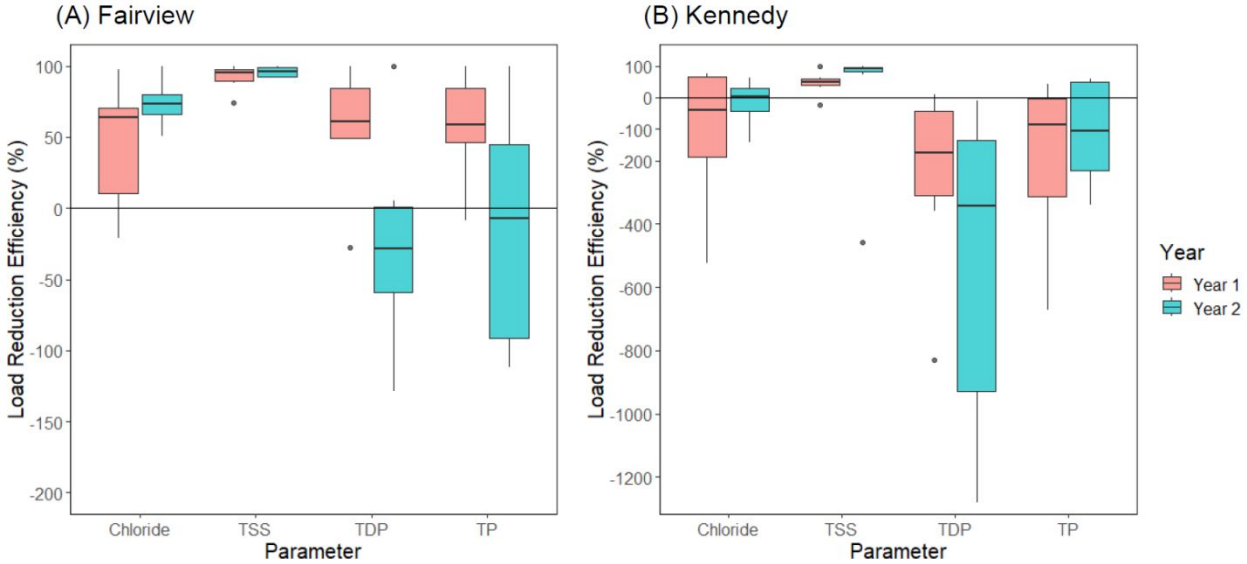

Figure 4. Boxplots for storm event load reduction efficiencies (%) observed in the field study for chloride, total suspended solids (TSS), total dissolved phosphorus (TDP), and total phosphorus (TP) at the (A) Fairview and (B) Kennedy sites in 2020 (year 1) and 2021 (year 2). Note the difference in y-axis scale for the two sites. Boxplots show median, interquartile range, range, and extreme values.

Table 1. Influent and effluent event mean concentrations (overall mean ±1 standard deviation) of chloride (Cl-), total suspended solids (TSS), particulate phosphorus (PP), total dissolved phosphorus (TDP), and total phosphorus (TP) during the two field study years at the Fairview and Kennedy subsurface gravel wetland sites. n = number of storms with concentration measurements.

Year	Site	Location	n	Cl^- (mg L^{-1})	TSS (mg L^{-1})	PP (mg L^{-1})	TDP (mg L^{-1})	TP (mg L^{-1})
2020	Fairview	Influent	6	117 ± 62	33 ± 26	0.087 ± 0.082	0.062 ± 0.025	0.14 ± 0.102
2020	Fairview	Effluent	4	128 ± 15	2.5 ± 0.5	0.019 ± 0.009	0.069 ± 0.012	0.089 ± 0.008
2021	Fairview	Influent	6	117 ± 13	51 ± 24	0.048 ± 0.37	0.061 ± 0.105	0.105 ± 0.141
2021	Fairview	Effluent	5	171 ± 99	7.5 ± 3.2	0.085 ± 0.054	0.145 ± 0.08	0.23 ± 0.128
2020	Kennedy	Influent	12	409 ± 471	19 ± 35	0.03 ± 0.035	0.03 ± 0.02	0.06 ± 0.027
2020	Kennedy	Effluent	11	566 ± 335	4.8 ± 1.6	0.07 ± 0.066	0.084 ± 0.04	0.16 ± 0.057
2021	Kennedy	Influent	8	306 ± 257	113 ± 164	0.065 ± 0.068	0.021 ± 0.013	0.127 ± 0.112
2021	Kennedy	Effluent	8	332 ± 238	14 ± 20	0.04 ± 0.029	0.105 ± 0.036	0.149 ± 0.045

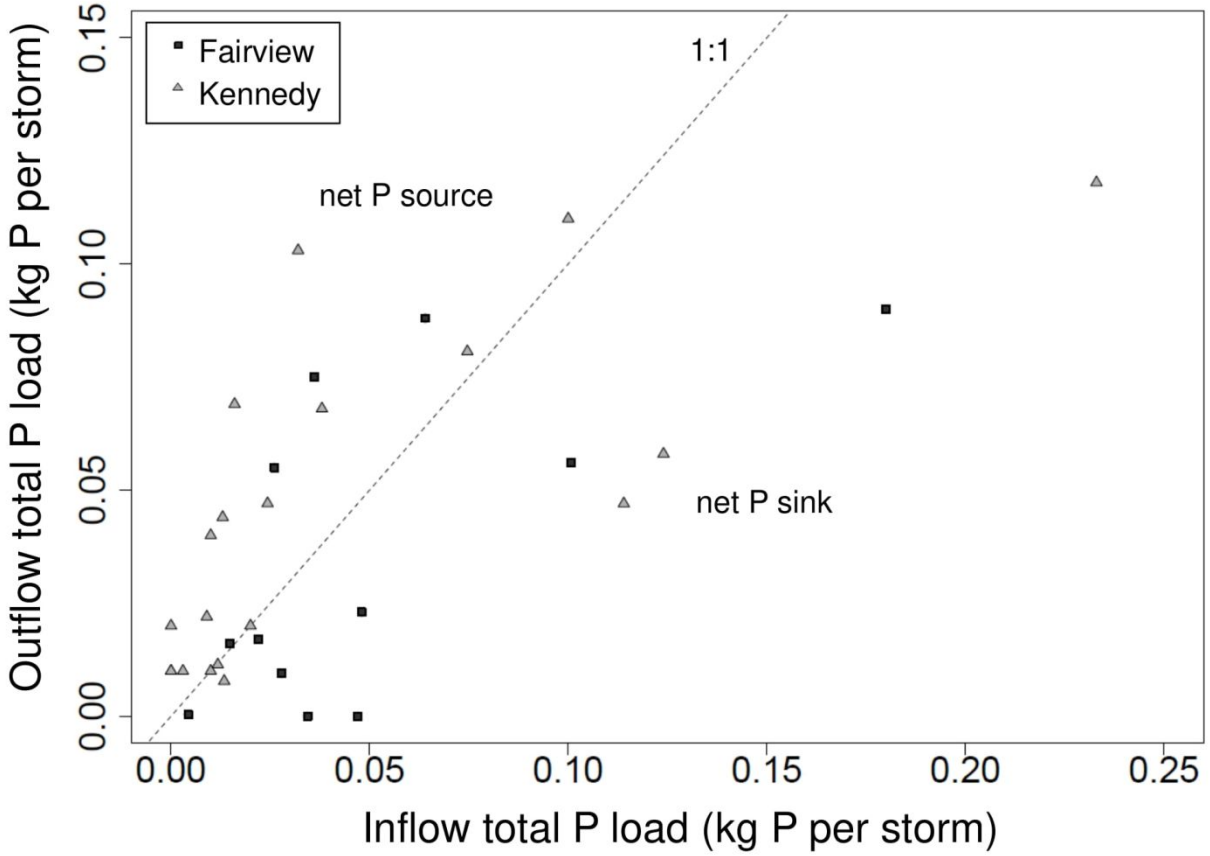

Figure 5. Outflow total phosphorus load (kg P) versus inflow total phosphorus load (kg P). Each point represents an individual storm at one of the two field sites. The 1:1 dashed line separates observations of the subsurface gravel wetlands acting as net P sources (upper left) versus net P sinks (lower right).

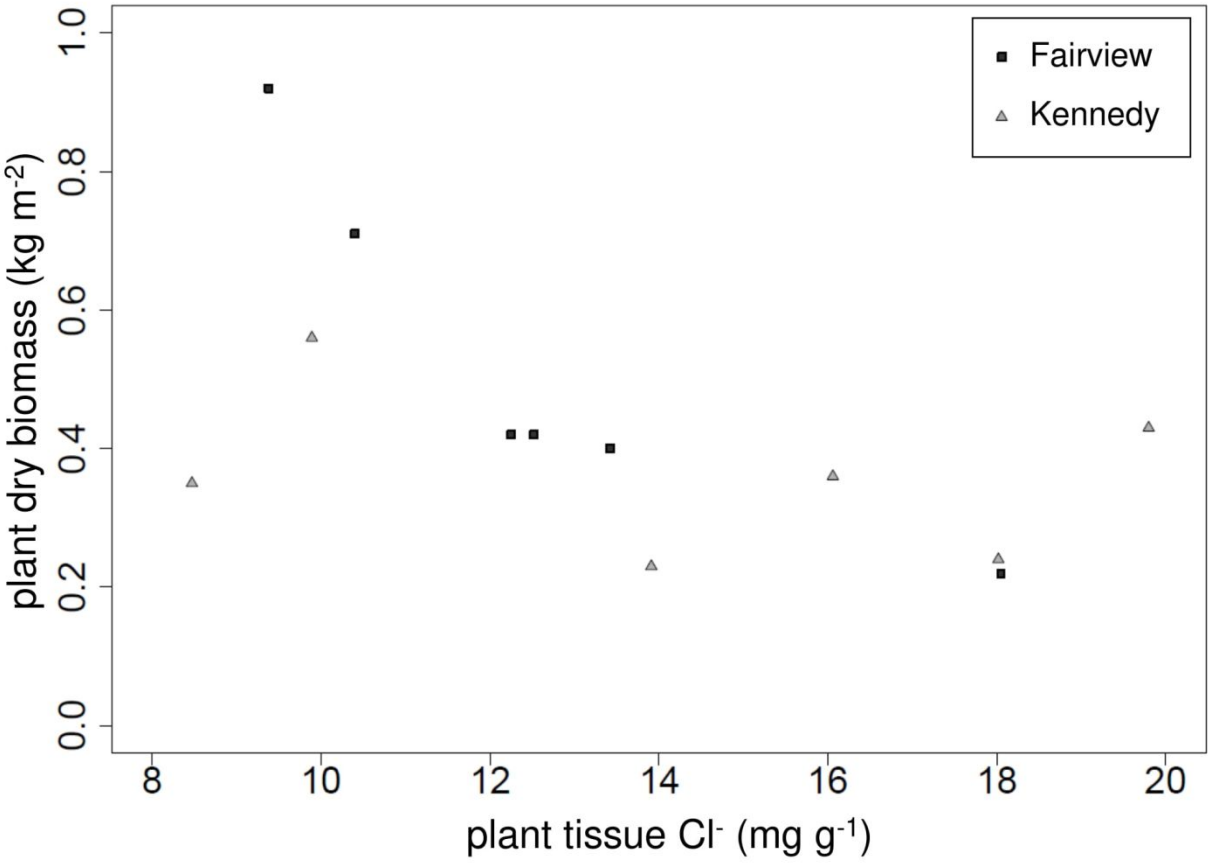

Figure 6. Aboveground plant biomass (dry kg m^{-2}) versus plant tissue chloride content (mg Cl^- g^{-1} dry) at the two field study sites (n = 3 samples per year x 2 years at each site).

Table 2. Gravel wetland soil chloride content and electrical conductivity (EC) at the two field sites in 2020 (n = 3 transect composite samples per site) and 2021 (n = 3 transect composite samples per site). Letters denote differences in a parameter at a given site between years (Welch Two Sample t-test, α = 0.05).

Field site	Fairview		Kennedy	
Year	soil Cl^- (mg kg^{-1} dry)	soil EC (mmhos cm^{-1})	soil Cl^- (mg kg^{-1} dry)	soil EC (mmhos cm^{-1})
2020	160 ± 97^a	0.30 ± 0.05^a	210 ± 35^a	0.29 ± 0.09^a
2021	333 ± 81^a	2.04 ± 0.39^b	208 ± 147^a	1.30 ± 0.29^b

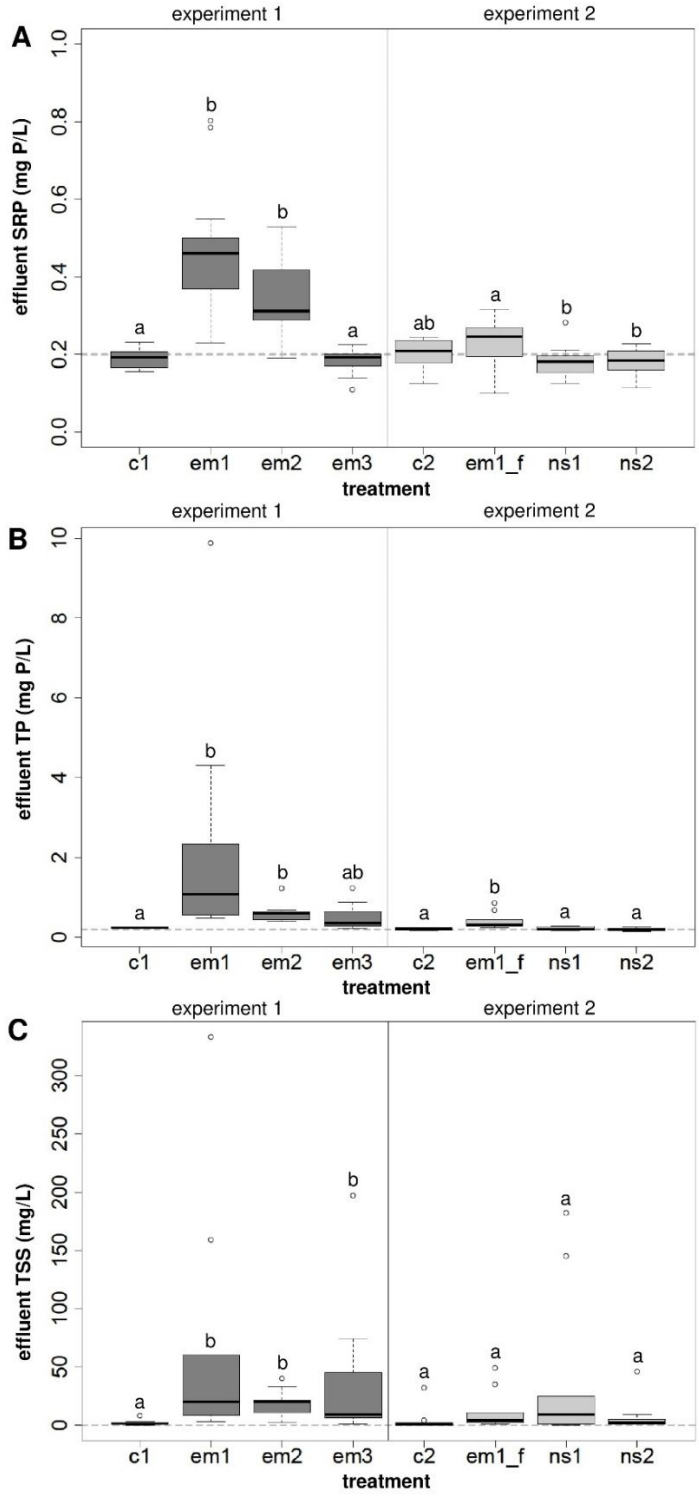

Figure 7. Boxplots for effluent (A) soluble reactive phosphorus (SRP), (B) total phosphorus (TP), and (C) total suspended solids (TSS) during the laboratory soil column experiments. Control cores (no soil) are denoted as c1 and c2. Soils included three engineered materials (em1, em2, em3), an engineered material after ~1 year in the field (em1_f), and two native soils from the field study sites (ns1, ns2). Two experiments were performed separately. Different letters denote significant differences (Dunn-Bonferroni, $p < 0.05$) between treatments for a given parameter within a given experiment.

Table 3. Physical and chemical characteristics for the six soil materials tested in the laboratory study, including three engineered materials (em1, em2, em3), an engineered material after ~1 year in the field (em1_f), and two native soils from the field study sites (ns1, ns2).

Sample	TS	VS	BD	pH	EC	TC	TN	Ksat
	% as-is basis	% dry basis	kg m^{-3}		mmhos cm^{-1}	% dry basis	% dry basis	cm hr^{-1}
em1	65.7	11.4	907.7	7.9	3.1	4.9	0.30	0.62 ± 0.37
em1_f	66.4	8.0	943.3	7.7	1.6	4.0	0.24	0.41 ± 0.14
em2	80.3	17.7	771.3	6.9	6.5	8.7	0.65	82.70 ± 33.26
em3	77.9	6.7	978.9	6.6	2.1	2.6	0.22	8.53 ± 7.34
ns1	70.9	9.7	812.8	5.3	0.1	4.0	0.23	7.07 ± 6.29
ns2	65.7	9.8	1014.5	6.3	0.1	3.5	0.31	3.05 ± 3.28

Table 4. Phosphorus metrics for the six soil materials tested in the laboratory study, including three engineered materials (em1, em2, em3), an engineered material after ~1 year in the field (em1_f), and two native soils from the field study sites (ns1, ns2). TP = total phosphorus, WEP = water-extractable phosphorus, MM-P = Modified Morgan phosphorus. Mehlich-3 and oxalate indicate to different extractions used to determine the phosphorus saturation ratio (PSR).

Sample	TP	WEP	MM-P	Mehlich-3				Oxalate			
	% dry basis	mg P kg^{-1}	mg P kg^{-1}	mg P kg^{-1}	mg Fe kg^{-1}	mg Al kg^{-1}	PSR_{M3}	mg P kg^{-1}	mg Fe kg^{-1}	mg Al kg^{-1}	PSR_{OX}
em1	0.16	41 ± 3	307	339	206	120	1.34	2055	1952	1904	0.63
em1_f	0.14	22 ± 0	192	316	294	228	0.75	586	1987	525	0.34
em2	0.37	27 ± 5	572	676	218	485	1.00	592	1206	431	0.51
em3	0.05	3 ± 2	30	161	123	1371	0.10	301	1261	2289	0.09
ns1	0.03	1 ± 1	2	10	344	1256	0.01	124	3646	1795	0.03
ns2	0.08	2 ± 0	3	56	337	1080	0.04	344	4251	1873	0.08

Table 5. Effluent characteristics of vertical upflow gravel columns in the laboratory experiment for soluble reactive phosphorus (SRP) (n = 6 days x 3 replicates per gravel), as well as total phosphorus (TP) and total suspended solids (TSS) (n = 1 composite x 3 replicates for each gravel). Asterisks denote significant differences (Dunn-Bonferroni or Tukey HSD, $p < 0.05$) between a gravel treatment and the control for a given parameter.

	effluent characteristics (mean ± 1 standard deviation, mg L^{-1})		
gravel	SRP	TP	TSS
control (none)	0.24 ± 0.10	0.41 ± 0.07	5 ± 2
granite	0.13 ± 0.06*	0.30 ± 0.03	15 ± 17
quartzite	0.14 ± 0.07	0.34 ± 0.08	4 ± 1
limestone	0.09 ± 0.06*	0.25 ± 0.04*	9 ± 8

Table 6. Laboratory chloride bioassay results using irrigation water containing 0%, 50%, and 100% of 650 mg Cl^- L^{-1}. Different letters denote significant differences (Dunn-Bonferroni or Tukey HSD, $p < 0.05$) across treatments for a given species-parameter combination (n = 5 each).

	Plant species			
	Broad-fruited bur-reed (*Sparganium eurycarpum*)		Shallow sedge (*Carex lurida*)	
Treatment	Biomass (dry g $plant^{-1}$)	Plant tissue Cl^- (mg Cl^- g^{-1} dry)	Biomass (dry g $plant^{-1}$)	Plant tissue Cl^- (mg Cl^- g^{-1} dry)
0 mg Cl^- L^{-1}	1.44 ± 0.29^a	2.8 ± 0.8^a	0.89 ± 0.61^a	19.0 ± 3.5^a
325 mg Cl^- L^{-1}	1.12 ± 0.18^a	17.6 ± 2.9^b	0.59 ± 0.18^{ab}	28.6 ± 5.6^{ab}
650 mg Cl^- L^{-1}	1.45 ± 0.28^a	33.8 ± 6.1^c	0.34 ± 0.11^b	36.4 ± 9.4^b