Stream restoration and sewers impact sources and fluxes of water, carbon, and nutrients in urban watersheds

Michael J. Pennino1, Sujay S. Kaushal2, Paul M. Mayer1, Ryan M. Utz3, and Curtis A. Cooper4

1US EPA, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Western Ecology Division, Corvallis, OR, USA
2University of Maryland, Department of Geology and Earth Systems Science Interdisciplinary Center, College Park, MD, USA
3Chatham University, Falk School of Sustainability, Pittsburgh, PA, USA
4Washington Department of Ecology, Environmental Assessment Program, Olympia, WA, USA

Correspondence to: Michael J. Pennino (michael.pennino@gmail.com)

Received: 15 October 2015 – Published in Hydrol. Earth Syst. Sci. Discuss.: 16 December 2015
Revised: 18 May 2016 – Accepted: 26 July 2016 – Published: 26 August 2016

Abstract. An improved understanding of sources and timing of water, carbon, and nutrient fluxes associated with urban infrastructure and stream restoration is critical for guiding effective watershed management globally. We investigated how sources, fluxes, and flowpaths of water, carbon (C), nitrogen (N), and phosphorus (P) shift in response to differences in urban stream restoration and sewer infrastructure. We compared an urban restored stream with two urban degraded streams draining varying levels of urban development and one stream with upland stormwater management systems over a 3-year period. We found that there was significantly decreased peak discharge in response to precipitation events following stream restoration. Similarly, we found that the restored stream showed significantly lower (p < 0.05) monthly peak runoff (9.4 ± 1.0 mm day−1) compared with two urban degraded streams (ranging from 44.9 ± 4.5 to 55.4 ± 5.8 mm day−1) draining higher impervious surface cover, and the stream-draining stormwater management systems and less impervious surface cover in its watershed (13.2 ± 1.9 mm day−1). The restored stream exported most carbon, nitrogen, and phosphorus at relatively lower streamflow than the two more urban catchments, which exported most carbon and nutrients at higher streamflow. Annual exports of total carbon (6.6 ± 0.5 kg ha−1 yr−1), total nitrogen (4.5 ± 0.3 kg ha−1 yr−1), and total phosphorus (161 ± 15 kg ha−1 yr−1) were significantly lower in the restored stream compared to both urban degraded streams (p < 0.05), but statistically similar to the stream draining stormwater management systems, for N exports. However, nitrate isotope data suggested that 55 ± 1% of the nitrate in the urban restored stream was derived from leaky sanitary sewers (during baseflow), statistically similar to the urban degraded streams. These isotopic results as well as additional tracers, including fluoride (added to drinking water) and iodide (contained in dietary salt), suggested that groundwater contamination was a major source of urban nutrient fluxes, which has been less considered compared to upland sources. Overall, leaking sewer pipes are a problem globally and our results suggest that combining stream restoration with restoration of aging sewer pipes can be critical to more effectively minimizing urban nonpoint nutrient sources. The sources, fluxes, and flowpaths of groundwater should be prioritized in management efforts to improve stream restoration by locating hydrologic hot spots where stream restoration is most likely to succeed.

1 Introduction

Urbanization significantly increases impervious surface cover (ISC), alters hydrologic regimes, and contributes to elevated organic carbon and nutrient exports in streams and rivers (e.g., Kaushal and Belt, 2012; Paul and Meyer, 2001; Walsh et al., 2005b). The growing impacts of urbanization on watershed nutrient exports have contributed to coastal eutrophication and hypoxia both regionally and globally.
(Nixon et al., 1996; Petrone, 2010). However, urban watersheds can differ significantly in carbon and nutrient sources and fluxes, and there are major questions regarding the potential influence of stream restoration and sewer infrastructure on the sources and fluxes of nutrients (e.g., Bernhardt et al., 2005; McMillan and Vidon, 2014; Passeport et al., 2013). Here, we characterize changes in streamflow variability pre- and post-restoration in an urban stream. We also compare sources and timing of fluxes of water, carbon, and nutrients in the urban restored stream with several urban degraded streams of varying levels of upland stormwater management and impervious surface cover.

The potential for increasing urbanization and climate change to alter hydrology and nutrient fluxes is a problem for cities globally (Julian and Gardner, 2014; Kaushal et al., 2014b; Old et al., 2006; Smith and Smith, 2015; Walsh et al., 2005a). It is well known that hydrologically connected impervious surfaces in urban watersheds create hydrologic regimes characterized by flow events with higher peaks, quicker time to peak, and shorter falling limbs – hereafter referred to as a “flashy” system (Konrad et al., 2005; Loperfido et al., 2014; Meierdiercks et al., 2010; Smith et al., 2013; Suduth et al., 2011b; Walsh et al., 2005b). Yet, little is known regarding the influence of stream restoration on hydrologic flashiness. Also, more work is necessary to characterize variability in fluxes of carbon and nutrients among urban watersheds, particularly for pulses (large changes in concentrations and fluxes over relatively short timescales) in urban restored and degraded streams (Kaushal et al., 2014b). Previous work indicates that pulses in carbon and nutrient exports can be influenced by the degree of hydrologic connectivity with impervious surfaces, sewer and stormwater infrastructure, and stream restoration features (e.g., Kaushal et al., 2014a; Newcomer et al., 2014). A recent global review and synthesis suggests that certain forms of stream restoration have potential to retain watershed nutrient exports particularly during baseflow, but further evaluation across streamflow is necessary (Newcomer-Johnson et al., 2016). Although stream restoration research is growing, the effects of stream restoration on minimizing pulses of water, carbon, and nutrient exports is still not clearly understood (Filoso and Palmer, 2011; Harrison et al., 2014; Newcomer et al., 2014).

One key to improved management of urban watersheds is a better understanding of contaminant sources and how they can shift across hydrologic variability in restored and urban degraded streams. Knowledge of the sources of chemical fluxes in urban restored streams is particularly lacking, even though stream restoration is currently a billion-dollar industry in the US (Bernhardt et al., 2005). In order to characterize contaminant sources, various biogeochemical and hydrologic tracers have been employed in other urban degraded watersheds. For example, recent studies have utilized N and O stable isotopes to determine sources of $\text{NO}_3^-$ (e.g., wastewater, atmospheric, or nitrification) (Burns et al., 2009; Kaushal et al., 2011; Kendall et al., 2007) in urban streams and rivers.

Tracking $\text{NO}_3^-$ can be improved when used in conjunction with additional tracers such as anions like fluoride and iodide (Kaushal et al., 2014a), where fluoride is applied as an additive to drinking water (Dean et al., 1950) and iodide is used in table salt (Waszkowiak and Szymandera-Buszk, 2008); therefore, their presence in streams may be considered an indicator of contamination by wastewater. Others have used fluorescence spectroscopy to determine dissolved organic matter sources and quality (e.g., labile vs. recalcitrant) (Baker, 2001; Cory et al., 2010; Smith and Kaushal, 2015), and to trace wastewater sources. Finally, stable isotopes of water have been used to characterize groundwater vs. surface water flowpaths (Gat, 1996; Harris et al., 1999; Kendall and Coplen, 2001). These techniques and others have been used globally to detect the influence of leaky sewer infrastructure on water quality (Ekklesia et al., 2015; Hall et al., 2016; Risch et al., 2015; Tran et al., 2014; Wolf et al., 2012) and it has been shown that sewer leaks have impacts during baseflow and stormflow (Divers et al., 2013, 2014; Phillips and Chalmers, 2009; Rose, 2007). The present study is unique in that it uses multiple tracers of contaminants to assess the effects of hydrologic variability on sources and fluxes of carbon and nutrients.

The objectives of this study were to characterize sources and timing of water, carbon, and nutrient fluxes in four urban watersheds with varying urban development and water management, including one site with extensive stream restoration. Our first objective was to compare the hydrologic response of the restored stream to precipitation events pre- and post-restoration. We predicted that the stream restoration, which reconnected the stream with its floodplain, had the potential to impact peak discharge and attenuate flashy flows. This would be due to the bankfull discharge overflowing onto the floodplain and infiltrating the floodplain soil and increasing groundwater contributions to baseflow (Bohnke et al., 2002; Cendon et al., 2010; Hester and Gooseff, 2010). In fact, floodplain reconnection is an a priori objective in restored streams in Baltimore, Maryland, USA and elsewhere globally (Banach et al., 2009; Duersken et al., 2005; Greenman-Pedersen Inc., 2003; Hoffmann et al., 2011; Klocker et al., 2009; Lamouroux et al., 2015). Our second objective was to compare metrics of hydrologic flashiness, and sources and timing of chemical fluxes in this restored urban stream with two urban degraded streams draining varying levels of urban development and one stream with extensive stormwater management systems in its catchment to assess the role of stream restoration and potential pollutant sources, such as leaky sanitary sewers. Research was conducted in watersheds that are part of the Baltimore Long-Term Ecological Research (BES LTER) project, which is described further below and elsewhere (www.beslter.org) (e.g., Groffman et al., 2004; Kaushal et al., 2011; Lindner and Miller, 2012; Meierdiercks et al., 2010; Ryan et al., 2010).
2 Methods

2.1 Site descriptions

All watersheds were located in the metropolitan region of Baltimore, Maryland, USA, in the Chesapeake Bay watershed (Fig. 1). Impervious surface cover (ISC) was calculated for each watershed using ArcGIS and based on averaging the ISC values obtained from the 2006 National Land Cover Database (NLCD), a 2 m satellite imagery obtained from the University of Vermont, and a roads and buildings polygon layer for Baltimore County. The amount of stormwater management (SWM) within each watershed was characterized using ArcGIS as the percentage of watershed drainage area that is managed by stormwater management systems. Data on the locations of SWM systems and the drainage area controlled by each SWM facility were provided by the Baltimore County, Maryland Department of Environmental Protection and Sustainability (BCMDEPS).

Four streams were chosen for this study: Minebank Run (MBR), an urban restored stream; Red Run (RRN), a stream with extensive upland SWM systems in its watershed; Dead Run (DRN), an urban degraded stream with upland SWM systems; and Powder Mill Run (PMR), an urban degraded stream with no SWM or stream restoration (Table 1). Details on the four watersheds in this study can be found in Table 1, with information on the % ISC, % SWM, median year built for development, range of flows, and range of flows sampled. DRN and PMR have the highest % ISC (45.7 % and 44.3 %, respectively), while MBR has intermediate % ISC (29.4 %), and RRN has the lowest % ISC (20.5 %) (Table 1). RRN has approximately 40 % SWM and DRN has 33 % SWM, while MBR and PMR have minimal SWM in their watersheds (Table 1).

About 95 % of Minebank Run’s mainstem has been restored (~ 5700 linear meters were restored, BCMDEPS); the headwaters were restored in 1998–1999 and the lower portions (directly above and below the stream gauge) were restored in 2004–2005. Restoration features at MBR include oxbows, redesigned channels, armoring, low connected floodplains, increased sinuosity, and step pools (Harrison et al., 2011; Kaushal et al., 2008b). The stormwater management at RRN is primarily in the lower portion of the watershed and includes detention ponds, wet ponds, bioretention, and sand filters, with its headwaters containing a quarry and low-density development on septic systems (BCMDEPS). DRN has stormwater management mainly in a portion of its headwaters, with primarily detention ponds (Fig. 1, Table 1, Smith et al., 2015). Also, RRN and MBR have broader undeveloped downstream riparian zones than either DRN or PMR.

Discharge was measured continuously at all of the four study watersheds: Minebank Run, Powder Mill Run, and Dead Run are gauged by the US Geological Survey (USGS gage numbers 0158397967, 01589305, and 01589330, respectively), while Red Run is gauged by the University of Maryland, Baltimore County Center for Urban Environmental Research and Education. Further details on stream site characteristics and the methods described below are in the Supplement.
In order to examine the hydrologic response of an urban stream to restoration, the relationship between effective precipitation (Ppt) and effective peak discharge ($Q_{pk}$) was estimated for Minebank Run pre- and post-restoration from 2001 to 2008. Discharge and precipitation data used in this analysis were from the US Geological Survey (USGS) National Water Information System.

There were 195 pre-restoration and 221 post-restoration dates used in the effective Ppt–$Q_{pk}$ analysis (where the designation of effective is used to specifically identify data that meet the assumptions of a measurable mechanism between precipitation leading to a discharge response). Regression lines were created in Minitab (release 14.2, Minitab, Inc. State College, PA, USA) to compare the precipitation amount (mm day$^{-1}$) with its associated daily peak discharge (cms) for the pre-restoration and the post-restoration data. Slope and intercept of these developed regression lines were compared using a general linear model in Minitab (ID 1248). See the Supplement for further details.

2.3 Water quality sampling and analyses

Water samples were collected at the MBR, RRN, DRN, and PMR stream gauge locations every 2–4 weeks (called “routinely sampled” water quality data from this point on) for 3 years (2010–2012) and longitudinally at 8–12 sampling points (300–1000 m apart) from the mouth to headwaters of each stream network during four different seasons: two winters (January 2010 and December 2010), one spring (April 2010), and one summer (June 2011). Samples were analyzed for total organic C (TOC), dissolved organic C (DOC), total Kjeldahl nitrogen (TKN), nitrate plus nitrite (NO$_3^-$ + NO$_2^-$), total phosphorus (TP), orthophosphate (PO$_4^{3-}$), iodide (I$^-$), fluoride (F$^-$), stable water isotopes ($\delta^{18}$H$_2$O and $\delta^{18}$O-H$_2$O, details below), C quality characterization (described further below), and NO$_3^-$ stable isotopes ($\delta^{15}$N-NO$_3^-$ and $\delta^{18}$O-NO$_3^-$, details below). All samples were analyzed using standard Environmental Protection Agency (EPA) methods by the US EPA National Risk Management Research Laboratory in Ada, Oklahoma, USA.

2.4 Nitrate and water stable isotope analyses and mixing models

Surface samples for $\delta^{15}$N-NO$_3^-$ and $\delta^{18}$O-NO$_3^-$ isotopes of dissolved NO$_3^-$ were filtered (0.45 µm), frozen, and shipped to the UC Davis Stable Isotope Facility (SIF) for analysis. The isotope composition of nitrate was measured following the denitrifier method (Casciotti et al., 2002; Sigman et al., 2001). Briefly, denitrifying bacteria were used to convert nitrate in water samples to N$_2$O gas, which was then analyzed by a mass spectrometer for stable isotopic ratios of N and O of nitrate ($^{15}$N/$^{14}$N and $^{18}$O/$^{16}$O). Values for $\delta^{15}$N-NO$_3^-$ and $\delta^{18}$O-NO$_3^-$ are reported as per mil (‰) relative to atmospheric N$_2$ ($\delta^{15}$N) or VSMOW ($\delta^{18}$O), according to $\delta^{15}$N or $\delta^{18}$O (‰) = [(R)sample / (R)standard − 1] × 1000, where R denotes the ratio of the heavy to light isotope ($^{15}$N/$^{14}$N or $^{18}$O/$^{16}$O). For data correction and calibration, UC Davis SIF uses calibration nitrate standards (USGS 32, USGS 34, and USGS 35) supplied by NIST (National Institute of Standards and Technology, Gaithersburg, MD). The long-term standard deviation for nitrate isotope samples at UC Davis SIF is 0.4‰ for $\delta^{15}$N-NO$_3^-$ and 0.5‰ for $\delta^{18}$O-NO$_3^-$. Previous studies (Kaushal et al., 2011; Kendall et al., 2007) indicate that the relative amounts of $\delta^{15}$N-NO$_3^-$ and $\delta^{18}$O-NO$_3^-$ can be used to determine specific sources of nitrate (i.e., fertilizer, atmospheric, or sewage-derived nitrate).

Stable nitrate isotope data were used to create a three-endmember isotope mixing model to determine the percent con-
tribution of different potential nitrate sources: wastewater, nitrification, or atmospheric-derived nitrate (Kaushal et al., 2011; Phillips, 2001), where

\[
f_{\text{wastewater}} = \frac{(\delta^{15}N_{\text{N}} - \delta^{15}N_{\text{A}})(\delta^{18}O_{\text{N}} - \delta^{18}O_{\text{A}}) - (\delta^{15}O_{\text{N}} - \delta^{15}O_{\text{A}})(\delta^{18}N_{\text{N}} - \delta^{15}N_{\text{A}})}{\delta^{18}O_{\text{A}} - \delta^{18}O_{\text{N}}}
\]

\[
f_{\text{atmospheric}} = \frac{(\delta^{15}N_{\text{S}} - \delta^{15}N_{\text{N}})(\delta^{18}O_{\text{W}} - \delta^{18}O_{\text{N}}) \times f_{\text{wastewater}}}{\delta^{18}O_{\text{A}} - \delta^{18}O_{\text{N}}}
\]

\[
f_{\text{nitrification}} = 1 - f_{\text{wastewater}} - f_{\text{atmospheric}},
\]

and \( f_{\text{wastewater}}, f_{\text{atmospheric}}, \) and \( f_{\text{nitrification}} \) are the fractions of nitrate from wastewater, atmospheric, or nitrification sources, respectively (also equivalent to % wastewater \( NO_3^-, \% \text{ atmospheric } NO_3^-, \) and % nitification \( NO_3^- \); \( \delta^{15}N_{\text{S}} \) or \( \delta^{18}O_{\text{S}} \) is the value (‰) for the nitrate sample; \( \delta^{15}N_{\text{N}} \) or \( \delta^{18}O_{\text{N}} \) is the end-member value (‰) for nitrification; \( \delta^{15}N_{\text{A}} \) or \( \delta^{18}O_{\text{A}} \) is the end-member value (‰) for atmospheric nitrate; and \( \delta^{15}N_{\text{W}} \) or \( \delta^{18}O_{\text{W}} \) is the end-member value (‰) for wastewater nitrite. End-member values for \( \delta^{15}N_{\text{NO}_3^-} \) and \( \delta^{18}O_{\text{NO}_3^-} \) for nitrification (−3 and 0, respectively) and atmospheric nitrate (−0.2 and 80, respectively) were obtained from an average of the values in Kendall et al. (2007). The wastewater \( \delta^{15}N_{\text{NO}_3^-} \) and \( \delta^{18}O_{\text{NO}_3^-} \) end-member value (35.4 and 13.3, respectively) was based on averaging the effluent nitrate isotope values measured from the Blue Plains waste water treatment plant in Washington D.C. (for monthly samples collected 2010−2011).

Water isotope (\( \delta^{2}H_{-H_2O} \) and \( \delta^{18}O_{-H_2O} \)) samples were collected from August 2010 to October 2011 and analyzed using a high temperature conversion elemental analyzer (TC/EA), a continuous flow unit, and an isotope ratio spectrometer (IRMS). A two-end-member mixing model (Buda and DeWalle, 2009; Kaushal et al., 2011; Williard et al., 2001) was created using \( \delta^{18}O_{\text{H}_2\text{O}} \) to distinguish between groundwater and atmospheric water sources, where

\[
\% \text{ groundwater} = \frac{\delta^{18}O_{\text{G}} - \delta^{18}O_{\text{R}}}{\delta^{18}O_{\text{G}} - \delta^{18}O_{\text{R}}} \times 100,
\]

and \% rainwater = 100 − \% groundwater, \( \delta^{18}O_{\text{G}} \) is the value (‰) for the streamwater sample, \( \delta^{18}O_{\text{R}} \) is the end-member value (‰) for rainwater, and \( \delta^{18}O_{\text{G}} \) is the end-member value (‰) for groundwater. End-member values for \( \delta^{2}H_{-H_2O} \) and \( \delta^{18}O_{-H_2O} \) from rainwater (−22.41 and −5.23, respectively) and groundwater (−44.02, and −7.995, respectively) were obtained from Kendall and Coplen (2001).

2.5 Fluorescence analyses for dissolved organic matter characterization

The lability (e.g., protein or humic-like) and sources (e.g., allochthonous or autochthonous) of dissolved organic mat-

ter were characterized using fluorescence excitation emission matrices (EEMs) (Cory and McKnight, 2005; Cory et al., 2010), using a Fluoromax-4 spectrofluorometer (Horiba, Jobin Yvon). Water samples were analyzed with an excitation range of 240−450 nm at 10 nm increments and an emission range of 290−600 nm at 2 nm increments. Fluorescence EEMs were instrument corrected, blank subtracted, and normalized by the water Raman signal following Cory et al. (2010). Standard inner-filter corrections (IFC) were not performed on samples because absorbance measurements were not attained for most samples (however, for a subset of samples, absorbance was collected using a scanning spectrophotometer, the inner-filter corrections were done, and it was found that there is <5 % difference in the EEM metric results, with and without IFC). We analyzed fluorescence EEMs for the following indices: fluorescence index, FI (McKnight et al., 2001); humification index, HIX (Huguet et al., 2009; Zsolnay et al., 1999); biological freshness index, BIX (Huguet et al., 2009); and protein-to-humic fluorescence intensities ratio, \( P/H \) ratio (Coble, 1996; Stolpe et al., 2010).

2.6 Estimation of annual watershed carbon, nutrient, and anion exports

Routine sampled concentration data, mean daily discharge, and the USGS FORTRAN program LOADEST (Runkel et al., 2004) were used to calculate the annual exports of all stream chemistry variables at each site. For clarification, the term load is used when referring to mass per the amount of time, while exports is used when referring to loads normalized by watershed area. Various methods have been employed for estimating annual nutrient exports (e.g., Cohn, 1995; Schwartz and Naiman, 1999). However, we chose LOADEST because it uses a multiple parameter regression model that accounts for bias, data censoring, and non-normality to minimize difficulties in load estimation (Qian et al., 2007). LOADEST uses three different statistical approaches to estimate load: adjusted maximum likelihood estimation (AMLE), maximum likelihood estimation (MLE), and least absolute deviation (LAD). As suggested by Runkel et al. (2004), AMLE was chosen when the calibration model errors (residuals) were normally distributed, while LAD was chosen when residuals were not normally distributed. LOADEST produced load estimates for daily nutrient loads and annual exports were calculated by summing daily load for each year and dividing by watershed area. Through analyses of model residuals and a comparison of the observed and estimated loads, none of the constituents where found to have bias in the LOADEST output (Runkel, 2013). Based on the mean daily runoff and estimated daily loads, flow duration and nutrient duration curves were quantified for each stream similar to previous studies (Duan et al., 2012; Shields et al., 2008; Sivirichi et al., 2011). Following Shields et al. (2008), we also calculated the F75 metric for each nutrient export, which is the runoff at which 75 % of
each nutrient is exported annually. Additionally, 95% confidence intervals were estimated for annual exports using a simplified bootstrap resampling approach similar to Efron and Tibshirani (1986) and Rustomji and Wilkinson (2008).

Samples were collected over a range of streamflow conditions. However, the largest flows were not sampled due to adherence to a random sampling scheme and logistic feasibility (see Table 1, Fig. S2 in the Supplement). Flow duration records, based on mean daily flow for 2010–2012, show that the majority of samples were collected during low to intermediate flows (Fig. S2). As a result, the daily load estimates from LOADEST may not accurately reflect flows higher than the highest flows sampled. Also, because mean daily discharge data were used instead of instantaneous discharge, there is likely increased uncertainty in the daily load estimates during storm event peak flow periods. However, because all four sites are within the same city and receive relatively the same rainfall during storm events, the relative annual loads estimated for the sites are comparable and it is appropriate to draw conclusions among the four study sites. Also, Carey et al. (2014) found no difference in annual load estimates in an urban watershed when using daily vs. instantaneous records of flow and nitrate concentration, though it was a significantly larger suburbanizing watershed. There are also likely differences in the effects of storms on carbon (C), nitrogen (N), and phosphorus (P) concentrations, since NO$_3^-$ is generally diluted during storms, whereas particulate organic nitrogen and P generally increases during storms (Bowes et al., 2005; Kaushal et al., 2008a). Additionally, when comparing the years sampled (2010–2012) to the full discharge record at each site (starting in 2001 for MBR, 2008 for RRN, 2005 for PMR, and 1998 for DRN), the range of streamflow during 2010–2012 contains 5 of the 10 highest flows recorded at all sites. Our sampling period also included streamflow equal to the lowest streamflow ever recorded at these gauges, indicating that 2010–2012 encompasses the full range of flows.

2.7 Characterizing hydrologic flashiness and pulses of C, N, and P exports

Metrics of hydrologic flashiness were calculated using daily and instantaneous discharge and precipitation data. Metrics consisted of the following variables: (1) average peak runoff, (2) hydrograph duration, (3) high-flow event frequency (monthly frequency of peaks above 3 × monthly median) (Utz et al., 2011), (4) mean monthly peak flow coefficient of variation, and (5) mean lag time (time between rainfall centroid and peak runoff) (Smith et al., 2013). Additionally, mean daily discharge data were used to calculate the flashiness index (average daily change in mean daily streamflow per month, divided by the mean monthly flow) (Poff et al., 2006a; Sudduth et al., 2011b), which is identical to the R-B index (Baker et al., 2004). Peak flow runoff is the only metric that accounts for watershed size. These metrics were chosen to provide a sense of how variability in urbanization affects typical stormflow characteristics and the variability in hydrologic response to storm events. Precipitation data used for lag-time calculations were 15 min interval rainfall data obtained from the National Atmospheric and Ocean Administration (NOAA) National Climatic Data Center (NOAA, 2014).

We also quantified the variability of routinely sampled carbon and nutrient source and concentration data and the daily export data from USGS LOADEST by calculating (1) the mean monthly coefficient of variation, (2) the mean difference (absolute value of change between consecutive daily exports or routinely sampled nutrient concentrations), and (3) the flashiness index (described above). These metrics were chosen to determine how differences in urbanization affect the variability or pulsing of C and nutrient sources, concentrations, and exports over time.

2.8 Statistical analyses

In order to compare all time series data (routinely sampled nutrient concentrations, stable isotopes, carbon quality indices, and monthly flashiness metrics at each stream site), we used a repeated measures ANOVA and post-hoc pairwise comparisons for each site with the Wilcox test (also called the Mann–Whitney test). This is a non-parametric rank sum test considered better suited for censored and skewed data (Cooper et al., 2014; Helsel and Hirsch, 1992; Lloyd et al., 2014). We used 95% confidence intervals for pairwise annual export comparisons. Analysis of covariance (ANCOVA) was performed to test for differences in regression slopes. Statistical analysis of trends were examined using Sen’s slope estimator and a Mann–Kendall test (Gilbert, 1987; Helsel and Hirsch, 2002). The Mann–Kendall test is a linear regression zero slope test of time-ordered data over time (Gilbert, 1987). Statistical analysis was performed using R software (R Development Core Team, 2013) or Minitab (Release 14.2, Minitab, Inc. State College, PA, USA), and MATLAB 8.1.0 (MATLAB and statistics toolbox release R2012a Student) was used for estimating hydrologic flashiness metrics in each stream for the period 2010–2012.

3 Results

3.1 Pre-restoration and post-restoration hydrologic analysis

Data from the analysis of the effective precipitation peak discharge relationship in MBR are shown in Fig. 2, for both the pre- and post-restoration periods (data during the restoration were not included in the analysis). The median storm depth was 7.6 mm during the pre-restoration period ($n = 195$) and 6.1 mm in the post-restoration period ($n = 221$). The median storm peak discharge was 0.7 cms in the pre-restoration period ($n = 195$) and 0.4 cms in the post-restoration period.
Storm peak discharge was 3.4 cms in the pre-restoration period. Associated with the 50 largest precipitation events, the median storm depth was 24.3 mm in the pre-restoration period (\(n = 50\text{largest}\)) and 22.4 mm in the post-restoration period (\(n = 50\text{largest}\)).

Regression lines and lines representing the 95% confidence bands were developed for both the pre- and post-restoration periods. The lower confidence band for the pre-restoration data is nearly identical to the upper confidence band for the post-restoration data. The pre-restoration line has a slope of 0.136 with an \(R^2\) of 0.74 (Eq. 5), whereas the post-restoration line has a slope of 0.117 with an \(R^2\) of 0.67 (Eq. 6) (Fig. 2).

\[
\text{pre - } Q_{\text{peak}} = -0.073 + 0.136 (\text{PPT}_{\text{pre}}) \\
\text{post - } Q_{\text{peak}} = -0.0596 + 0.117 (\text{PPT}_{\text{post}})
\]

Comparison of the slopes and intercepts of the above equations using a general linear model found that the intercepts were not significantly different but the slopes were significantly different (\(p = 0.019\)). Therefore, the different slopes indicate that regression lines are different between the pre- and post-restoration effective precipitation – effective peak discharge relationship.

### 3.2 Sources of water, carbon, and nitrogen exports among urban watersheds

 Routinely sampled stable deuterium (\(\delta^2\)H) and \(\delta^{18}\)O water isotopes were not significantly different between sites, including the restored stream. MBR (\(p > 0.05\)) (Table 2), and there was also no separation when plotting \(\delta^{18}\)O-H2O vs.

---

**Table 2.** Comparisons of water, carbon, and nitrate sources (mean ± SE) among the four urban watersheds.

<table>
<thead>
<tr>
<th>Water isotopes</th>
<th>MBR</th>
<th>RRN</th>
<th>PMR</th>
<th>DRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\delta^2)H-H2O</td>
<td>-43 ± 1.8(^a)</td>
<td>-44 ± 2.0(^a)</td>
<td>-43 ± 2.5(^a)</td>
<td>-43 ± 3.0(^a)</td>
</tr>
<tr>
<td>(\delta^{18})O-H2O</td>
<td>-6.7 ± 0.2(^a)</td>
<td>-6.9 ± 0.2(^a)</td>
<td>-6.6 ± 0.3(^a)</td>
<td>-6.6 ± 0.4(^a)</td>
</tr>
<tr>
<td>% Groundwater</td>
<td>50 ± 5(^a)</td>
<td>57 ± 6(^a)</td>
<td>47 ± 6(^a)</td>
<td>40 ± 7(^a)</td>
</tr>
<tr>
<td>% Rainwater</td>
<td>50 ± 5(^a)</td>
<td>43 ± 6(^a)</td>
<td>53 ± 6(^a)</td>
<td>60 ± 7(^a)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carbon quality</th>
<th>HIX</th>
<th>BIX</th>
<th>FI</th>
<th>(P/H) ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBR</td>
<td>0.87 ± 0.01(^a)</td>
<td>0.81 ± 0.02(^b)</td>
<td>0.80 ± 0.01(^c)</td>
<td>0.83 ± 0.02(^ab)</td>
</tr>
<tr>
<td>RRN</td>
<td>0.73 ± 0.04(^a)</td>
<td>0.64 ± 0.03(^a)</td>
<td>0.75 ± 0.04(^a)</td>
<td>0.78 ± 0.04(^a)</td>
</tr>
<tr>
<td>PMR</td>
<td>1.20 ± 0.05(^a)</td>
<td>1.15 ± 0.05(^bc)</td>
<td>1.16 ± 0.05(^ac)</td>
<td>1.26 ± 0.05(^ac)</td>
</tr>
<tr>
<td>DRN</td>
<td>0.73 ± 0.07(^ab)</td>
<td>0.66 ± 0.06(^a)</td>
<td>1.11 ± 0.10(^c)</td>
<td>0.89 ± 0.10(^b)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nitrate isotopes</th>
<th>(\delta^{15})N-NO(_3)</th>
<th>(\delta^{18})O-NO(_3)</th>
<th>% Wastewater</th>
<th>% Atmospheric</th>
<th>% Nitrification</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBR</td>
<td>7.0 ± 0.2(^ab)</td>
<td>5.0 ± 0.4(^a)</td>
<td>53 ± 1.0(^a)</td>
<td>8.7 ± 1.0(^ab)</td>
<td>38 ± 0.7(^a)</td>
</tr>
<tr>
<td>RRN</td>
<td>6.3 ± 0.2(^a)</td>
<td>4.0 ± 0.3(^a)</td>
<td>51 ± 1.1(^b)</td>
<td>7.6 ± 1.0(^ab)</td>
<td>41 ± 0.5(^b)</td>
</tr>
<tr>
<td>PMR</td>
<td>8.1 ± 0.2(^c)</td>
<td>5.9 ± 0.6(^a)</td>
<td>56 ± 1.2(^b)</td>
<td>9.4 ± 1.7(^ab)</td>
<td>34 ± 0.7(^c)</td>
</tr>
<tr>
<td>DRN</td>
<td>7.5 ± 0.2(^bc)</td>
<td>8.0 ± 0.9(^c)</td>
<td>52 ± 1.8(^ab)</td>
<td>15 ± 2.5(^b)</td>
<td>33 ± 0.9(^c)</td>
</tr>
</tbody>
</table>

MBR indicates Minebank Run, RRN indicates Red Run, PMR indicates Powder Mill Run, DRN indicates Dead Run. Different letters (\(^a\), \(^b\), \(^c\)) indicate significant differences (\(p < 0.05\)), based on pairwise comparisons of 3 years of routinely sampled data. HIX indicates humification index; BIX indicates biological freshness index; FI indicates fluorescence index; \(P/H\) ratio indicates protein-to-humic ratio.

---

**Figure 2.** Effective precipitation and effective discharge for Minebank Run. Best-fit regression lines and 95% confidence lines included.
Comparison of (a) water isotopes ($\delta^2$H-H$_2$O vs. $\delta^{18}$O-H$_2$O), (b) C quality metrics (biological freshness index vs. protein-to-humic ratio), and (c) nitrate isotopes ($\delta^{15}$N-NO$_3^-$ vs. $\delta^{18}$O-NO$_3^-$). GMWL indicates global meteoric water line, LMWL indicates local meteoric water line (Craig, 1961; Kendall and Coplen, 2001).

$\delta^2$H-H$_2$O (Fig. S4a). Water isotope mixing model results also indicate no difference in the percent contribution of groundwater or rainwater sources to the stream between sites (Table 2). However, longitudinal data indicate that watersheds with higher % ISC (PMR and DRN) had significantly higher ($p < 0.05$) $\delta^{18}$O-H$_2$O isotope values in the headwaters than RRN and higher $\delta^2$H-H$_2$O isotopes ($p = 0.03$ for PMR & $p = 0.057$ for DRN) in the headwaters than MBR (Fig. 3a) during one winter sampling, indicative of greater evaporation of surface water at the more urban streams.

Fluorescence analyses indicated that the watersheds with greater % ISC (PMR and DRN) transported more labile organic matter than the less urban site, RRN, as suggested by trends in the biological freshness index (BIX, $p < 0.05$) and protein-to-humic ($P/H$) ratio ($p < 0.05$, Fig. 3b, Table 2), while MBR, the restored stream, was not different than the more urban sites (Fig. 3b, Table 2). Only one of the more urban degraded streams (PMR) had greater $\delta^{15}$N-NO$_3^-$ and contributions of NO$_3^-$ from wastewater than the restored stream (MBR) and the stream in the least developed watershed with SWM systems (RRN, $p < 0.05$); the most urban stream (DRN) was not significantly different than the other streams (Fig. 3c, Table 2). The percent contribution of NO$_3^-$ from atmospheric sources, however, was greater in the watershed with the highest % ISC (DRN) compared to the watershed with the lowest % ISC (RRN) ($p < 0.05$, Table 2), but not different than the restored stream (MBR). Additionally, all sites showed a significant decline in $\delta^{15}$N-NO$_3^-$ with increasing runoff, and the two least urban sites (RRN and MBR), including the restored stream MBR, showed steeper slopes than PMR and DRN ($p < 0.05$,
Table 3. Annual runoff, C, N and P exports (mean ± 95 % confidence intervals) for the 2010, 2011, and 2012 calendar years.

<table>
<thead>
<tr>
<th>Year</th>
<th>MBR</th>
<th>RRN</th>
<th>PMR</th>
<th>DRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runoff (mm yr⁻¹)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>522 ± 72</td>
<td>325 ± 23</td>
<td>497 ± 83</td>
<td>625 ± 117</td>
</tr>
<tr>
<td>2011</td>
<td>647 ± 88</td>
<td>504 ± 114</td>
<td>639 ± 110</td>
<td>851 ± 176</td>
</tr>
<tr>
<td>2012</td>
<td>412 ± 75</td>
<td>382 ± 61</td>
<td>498 ± 105</td>
<td>564 ± 164</td>
</tr>
<tr>
<td>MEAN</td>
<td>527 ± 45</td>
<td>404 ± 44</td>
<td>545 ± 58</td>
<td>680 ± 89</td>
</tr>
</tbody>
</table>

| Carbon (kg ha⁻¹ yr⁻¹) |  |  |  |  |
| DOC  |  |  |  |  |
| 2010 | 6.7 ± 1.3 | 6.2 ± 0.8 | 15 ± 3 | 28 ± 7  |
| 2011 | 9.1 ± 1.6 | 22 ± 8 | 27 ± 5 | 57 ± 15 |
| 2012 | 5.7 ± 1.5 | 11 ± 3 | 17 ± 4 | 33 ± 12 |
| MEAN | 7.2 ± 1  | 13 ± 3  | 20 ± 2c | 39 ± 7d |
| TOC  |  |  |  |  |
| 2010 | NA | NA | NA | NA |
| 2011 | 8.1 ± 1.2 | 26 ± 11 | 40 ± 11 | 45 ± 11 |
| 2012 | 5.1 ± 1.1 | 14 ± 5 | 26 ± 9 | 30 ± 10 |
| MEANa | 6.6 ± 0.5 | 20 ± 4b | 33 ± 5c | 38 ± 5e |

| Nitrogen (kg ha⁻¹ yr⁻¹) |  |  |  |  |
| NO₃⁻ |  |  |  |  |
| 2010 | 4.1 ± 0.3 | 3.7 ± 0.2 | 6.6 ± 0.9 | 4.1 ± 0.6 |
| 2011 | 4.6 ± 0.4 | 4.1 ± 0.4 | 8.0 ± 1.1 | 5.3 ± 0.8 |
| 2012 | 2.9 ± 0.3 | 3.7 ± 0.2 | 6.3 ± 1.1 | 3.6 ± 0.7 |
| MEAN | 3.9 ± 0.2a | 3.8 ± 0.2a | 7.0 ± 0.6b | 4.3 ± 0.4a |
| TN  |  |  |  |  |
| 2010 | 4.8 ± 0.4 | 4.4 ± 0.3 | 9.1 ± 1.5 | 6.7 ± 1.2 |
| 2011 | 5.4 ± 0.5 | 5.4 ± 0.7 | 11.6 ± 2.1 | 8.8 ± 1.6 |
| 2012 | 3.4 ± 0.5 | 4.6 ± 0.7 | 9.1 ± 2.1 | 5.9 ± 1.6 |
| MEANa | 4.5 ± 0.3a | 4.8 ± 0.3a | 9.9 ± 1.1b | 7.1 ± 0.9c |

| Phosphorus (kg ha⁻¹ yr⁻¹) |  |  |  |  |
| PO₄⁻³ |  |  |  |  |
| 2010 | 60 ± 9 | 58 ± 6 | 134 ± 22 | 167 ± 37 |
| 2011 | 75 ± 11 | 120 ± 29 | 172 ± 30 | 255 ± 62 |
| 2012 | 47 ± 10 | 66 ± 11 | 134 ± 33 | 122 ± 40 |
| MEAN | 61 ± 6a | 81 ± 11b | 147 ± 17c | 181 ± 28c |
| TP |  |  |  |  |
| 2010 | 138 ± 19 | 160 ± 17 | 290 ± 51 | 330 ± 60 |
| 2011 | 202 ± 29 | 431 ± 136 | 379 ± 72 | 454 ± 92 |
| 2012 | 143 ± 30 | 314 ± 89 | 298 ± 66 | 306 ± 76 |
| MEANa | 161 ± 15a | 302 ± 54b | 322 ± 37b | 363 ± 45b |

| Wastewater indicator anions (g ha⁻¹ yr⁻¹) |  |  |  |  |
| F⁻ |  |  |  |  |
| 2010 | 230 ± 11 | b.d. | 2.1 × 4 ± 1.0 × 4 | 726 ± 87 |
| 2011 | 235 ± 10 | b.d. | 1.8 × 3 ± 5.5 × 3 | 606 ± 91 |
| 2012 | 67 ± 5 | b.d. | 5.4 × 3 ± 3.5 × 3 | 281 ± 45 |
| MEAN | 177 ± 5a | NA | 1.5 × 3 ± 4.0 × 3b | 583 ± 45c |
| I⁻ |  |  |  |  |
| 2010 | 19 ± 1 | 21 ± 1 | 20 ± 2 | 50 ± 7 |
| 2011 | 29 ± 2 | 41 ± 8 | 39 ± 4 | 85 ± 13 |
| 2012 | 16 ± 1 | 24 ± 4 | 22 ± 2 | 46 ± 8 |
| MEAN | 21 ± 1b | 29 ± 3b | 27 ± 2b | 60 ± 6a |

MBR indicates Minebank Run, RRN indicates Red Run, PMR indicates Powder Mill Run, DRN indicates Dead Run. Different letters (a, b, c, or d) indicate significant differences, based on 95 % CI of exports. DOC indicates dissolved organic C; TOC indicates total organic C; TN indicates total nitrogen; TP indicates total phosphorus; b.d. indicates values below detection. * Note that this range is from 2011–2012, unlike the others.
Fig. 4a). Also, the more urban sites (PMR and DRN) showed
pulses in $\delta^{18}$O-NO$_3^-$ during rain events (Fig. 4b), which sug-
gests that atmospheric NO$_3^-$ contributions can increase with run-
off.

Longitudinally, after a spring rain event, the wastewater
nitrate signal (based on $\delta^{15}$N-NO$_3^-$ values) declines from the
headwater to the mouth in the more urban watershed (DRN),
while the $\delta^{15}$N-NO$_3^-$ values are relatively constant at the
restored stream, MBR, and least urban watershed, RRN
(Fig. 5a). Conversely, during summer baseflow, the $\delta^{15}$N-
NO$_3^-$ values are relatively steady at all four sites, but with
the more urban streams (PMR and DRN) having consistently
higher $\delta^{15}$N-NO$_3^-$ values (Fig. 5b). The contribution of at-
mospheric nitrate (based on $\delta^{18}$O-NO$_3^-$ values) during the
spring high flow period generally increased downstream for
the more urban degraded streams, but decreased for the re-
stored stream, MBR, and stayed the same longitudinally for
the less urban watershed with SWM systems (RRN, Fig. 5c).
There was little difference in the $\delta^{18}$O-NO$_3^-$ values longitu-
dinally for summer (Fig. 5d).

### 3.3 Carbon, nutrient, and anion exports among urban watersheds

Among watersheds, annual DOC export showed up to a 5-
fold difference and there was up to a 2-fold difference in an-
nual TP exports. The most urban watershed, DRN, exhibited
the highest and the restored stream, MBR, exhibited the low-
est annual TOC and TP exports (Table 3, p < 0.05 for DRN
vs. MBR). The restored stream and the least urban stream
draining a watershed with SWM systems, RRN, also exhib-
itied lower annual total N (TN) exports compared to the more
urban catchments (DRN and PMR) (p < 0.05, Table 3). An-
nual NO$_3^-$ exports were not significantly different between
the restored stream and the most urban degraded stream,
DRN (Table 3). Annual exports of wastewater indicator an-
ions (fluoride and iodide) showed up to 3-fold differences among watersheds, with DRN exhibiting the highest and the
restored stream, MBR, the lowest annual exports (Table 3,
$p < 0.05$ for DRN vs. MBR).

### 3.4 Flashiness of water, carbon, and nutrient exports
among urban watersheds

The sites with greater % ISC (PMR and DRN) had signifi-
cantly higher monthly peak runoff, mean coefficient of vari-
ation of peak runoff, and flashiness index ($p < 0.05$, Table 4,
Fig. 6a) than RRN and the restored stream MBR. RRN (the
site with lowest % ISC) also had lower frequency of peak
flow runoff events above $3 \times$ median monthly runoff, and
longer hydrograph duration than the other sites (Table 4). Hy-
drologic lag time was not significantly different among sites
(Table 4).

The two most urban streams (PMR and DRN) showed
more variable and pulsed runoff and exports, based on the
the flashiness index (Fig. 6) and the time series of daily ex-
ports for C, N, and P (Fig. 7). Typically, exports of C, N,
P, and wastewater indicator anions (F$^-$ and I$^-$), showed a
lower flashiness index (less variable or pulsed) for sites with
lower % ISC including the restored stream (MBR and RRN;
Fig. 6b–d). Based on nutrient duration curves, the urban de-
graded sites with higher % ISC (PMR and DRN) exported
more C, N, and P during higher flows, while the restored
stream MBR and the less urban sites with SWM systems
Table 4. Hydrologic flashiness metrics (mean ± SE).

<table>
<thead>
<tr>
<th>Watershed area (km²)</th>
<th>% ISC</th>
<th>Mean peak flow runoff (mm day⁻¹)</th>
<th>Monthly CV (%) of peak runoff</th>
<th>Freq. peaks per month &gt; 3 × monthly median Q</th>
<th>Mean hydrograph duration (h)</th>
<th>Mean Lag time (h)</th>
<th>Avg. monthly flash index</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBR</td>
<td>5.3</td>
<td>21.9</td>
<td>9.4 ± 1.0ab</td>
<td>92 ± 6ab</td>
<td>5.7 ± 0.4abc</td>
<td>40 ± 1.7a</td>
<td>0.9 ± 0.1a</td>
</tr>
<tr>
<td>RRN</td>
<td>19.1</td>
<td>14.6</td>
<td>13.2 ± 1.9ab</td>
<td>63 ± 8ab</td>
<td>2.2 ± 0.3b</td>
<td>64 ± 2.4b</td>
<td>0.5 ± 0.1b</td>
</tr>
<tr>
<td>PMR</td>
<td>9.4</td>
<td>35.5</td>
<td>55.4 ± 5.8c</td>
<td>104 ± 7a</td>
<td>5.3 ± 0.5c</td>
<td>30 ± 1.4c</td>
<td>1.0 ± 0.1b</td>
</tr>
<tr>
<td>DRN</td>
<td>14.3</td>
<td>39.3</td>
<td>44.9 ± 4.5c</td>
<td>116 ± 7a</td>
<td>7.0 ± 0.5c</td>
<td>50 ± 1.5d</td>
<td>1.2 ± 0.1c</td>
</tr>
</tbody>
</table>

MBR indicates Minebank Run, RRN indicates Red Run, PMR indicates Powder Mill Run, DRN indicates Dead Run. Different letters (a, b, c, or d) indicate significant differences (p < 0.05) based on pairwise comparisons of 3 years of mean monthly flashiness metrics. ISC indicates impervious surface cover; CV indicates coefficient of variation; Q indicates discharge; Lag time indicates time between rainfall centroid and peak runoff; flash index indicates average daily change in mean daily streamflow per month, divided by the mean monthly flow.

Figure 6. Comparison of the flashiness index for (a) runoff, (b) dissolved organic carbon (DOC) concentration and export, (c) total nitrogen (TN) concentration and export, and (d) total phosphorus (TP) concentration and export. Conc. indicates concentration. Error bars are standard errors of the mean. N = 36, from averaging the monthly flashiness index over 3 years. Flashiness index indicates average change in daily export or routinely sampled concentration per month, divided by the mean monthly export or concentration per month.

(RRN) exported more during lower flows (Fig. 8). Similarly, the F75 metric showed that 75% of NO₃⁻, TN, PO₄³⁻, F⁻, and I⁻ export occurred for the site with restoration (MBR) and with lower % ISC and more SWM (RRN) typically at lower runoff than in higher % ISC sites PMR and DRN (Table 5).

4 Discussion

Our results show that watershed urbanization increases hydrologic flashiness and pulses in exports of carbon, nutrients, and atmospheric nitrate sources. From a management perspective, our results suggest that combining stream restoration with sewer infrastructure restoration has the potential to minimize sources, fluxes, and flowpaths of nutrients. Overall, impervious surface cover appeared to be an important indicator of timing of fluxes from the watersheds. Watersheds with older sewer infrastructure and higher ISC (DRN and PMR) showed significant differences in NO₃⁻ sources and C, N, and P exports than the stream restoration site (MBR) and the less urban stream with SWM systems in its catchment (RRN). Below, we discuss potential effects of stream restoration and sewer infrastructure on sources, fluxes, and flowpaths of nutrients across a broader range of sites and urban development.

4.1 Pre-restoration and post-restoration hydrologic analysis

Restoration had subtle but statistically significant impacts on hydrology by decreasing peak discharges during storm...
Figure 7. Routinely sampled (a) runoff, (b) DOC export, (c) NO$_3^-$ export, and (d) PO$_4^{3-}$ export over time.

Table 5. F75 metric: the runoff below which 75% of nutrients are exported.

<table>
<thead>
<tr>
<th>Site</th>
<th>F75 DOC (mm day$^{-1}$)</th>
<th>F75 TOC (mm day$^{-1}$)</th>
<th>F75 NO$_3^-$ (mm day$^{-1}$)</th>
<th>F75 TN (mm day$^{-1}$)</th>
<th>F75 PO$_4^{3-}$ (mm day$^{-1}$)</th>
<th>F75 TP (mm day$^{-1}$)</th>
<th>F75 I$^-$ (mm day$^{-1}$)</th>
<th>F75 F$^-$ (mm day$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBR</td>
<td>16.1</td>
<td>15.1</td>
<td>6.9</td>
<td>7.3</td>
<td>12.4</td>
<td>12.4</td>
<td>4.6</td>
<td>2.8</td>
</tr>
<tr>
<td>RRN</td>
<td>34.1</td>
<td>44.5</td>
<td>2.2</td>
<td>3.1</td>
<td>11.7</td>
<td>23.7</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>PMR</td>
<td>22.8</td>
<td>38.1</td>
<td>14.0</td>
<td>20.5</td>
<td>20.8</td>
<td>20.8</td>
<td>8.6</td>
<td>34.0</td>
</tr>
<tr>
<td>DRN</td>
<td>57.3</td>
<td>37.4</td>
<td>25.9</td>
<td>28.3</td>
<td>39.3</td>
<td>29.8</td>
<td>17.9</td>
<td>16.5</td>
</tr>
</tbody>
</table>

MBR indicates Minebank Run, RRN indicates Red Run, PMR indicates Powder Mill Run, DRN indicates Dead Run. DOC indicates dissolved organic C; TOC indicates total organic C; TN indicates total nitrogen; TP indicates total phosphorus. Similar to Shields et al. (2008).

Events in this flashy system. Stream restoration, which involved reconnection of the floodplain, was likely able to reduce peak discharge by increasing infiltration when bankfull discharge overflows onto the floodplain during storm events (Bohneke et al., 2002; Cendon et al., 2010; Hester and Goossens, 2010). In urban settings, imperious surfaces are identified as the primary mechanism for flashy hydrology and stream channel degradation (Doheny et al., 2006; Leopold, 1968; Paul and Meyer, 2001; Walsh et al., 2005b). As a result, small increases in impervious surface cover elicit disproportionately large reductions in water quality and biotic integrity (Brabec et al., 2002). Therefore, even small reductions in hydrologic flashiness may be an important benefit of restoration.

The Ppt–$Q_{pk}$ regression method for urban stream analysis used readily available data sources that are potentially applicable where there have been management changes but typical rainfall–runoff metrics do not apply (i.e., curve numbers). A clear understanding of statistically significant effects (i.e., decreased peak discharges) due to restoration are necessary to support decisions to enhance restoration beyond simple channel reconfigurations and make more active use of floodplains and/or synergistically integrating stormwater management in the uplands. The proposed Ppt–$Q_{pk}$ approach, however, does not quantify change, but only indicates if a change in the peak discharge has occurred. Also, this regression method may not be applicable to larger basins which have different routing pathways and processes that may not occur at the same rate as in a smaller basin (Ziemer and Lisle, 1998). Further study is needed to comprehensively evaluate the effects of stream restoration on hydrologic responses in larger basins and different climates. The wide availability of high-resolution precipitation data and discharge data make this a potentially useful method to evaluate management effects.
4.2 Sources of water, carbon, and nitrogen exports among urban watersheds

All four watersheds showed no significant differences in water isotope signatures, potentially due to complex mixing of surface water with groundwater and leaky urban water infrastructure, which is common among urban watersheds of the Baltimore LTER site (Kaushal and Belt, 2012; Kaushal et al., 2014a; Newcomer et al., 2014). Previous work has suggested that urban watersheds receive considerable inputs of water from a combination of groundwater and leaky urban water infrastructure (Bhaskar and Welty, 2012; Kaushal and Belt, 2012; Kaushal et al., 2014a). Recent evidence suggests that the urban stream corridor can be an important nonpoint source (or sink) of some pollutants due to leaky sewer infrastructure, groundwater contributions, and also in-stream production of labile organic carbon (Divers et al., 2013; Kaushal et al., 2014a; Newcomer et al., 2014). In fact, City of Baltimore has detailed records for the dates and locations of sewer overflows through their open data website (https://data.baltimorecity.gov/) and these sewer overflows have occurred within the watersheds of this study.

The more urbanized watersheds (PMR and DRN), as well as the restored stream, MBR, contained more labile dissolved organic matter than the more recently developed and less urban watershed with SWM systems (RRN). From studies throughout the globe, it is known that protein-like and more bioavailable or labile organic matter is typically associated with wastewater carbon sources (Baker, 2001; Goldman et al., 2012; Li et al., 2015; Yu et al., 2015). As a result, the higher BIX, P/H ratio, and protein-like organic matter in the restored stream MBR, as well as the more urban watersheds (PMR and DRN), is likely due to leaky sewers typically found in older urban watersheds (Hudson et al., 2008; Kaushal et al., 2011). The watershed restoration in this study are not influenced by combined sewer overflows or typical point source discharges of wastewater. However, the more labile organic matter found in the more urban streams may also be due to lack of a riparian zone, and more light availability, typical of degraded urban streams (Goetz et al., 2003), which promotes autotrophic growth and more biologically labile DOM (Huguet et al., 2009; McKnight et al., 2001; Pennino et al., 2014; Petrone et al., 2011). DOM derived from autotrophic production also tends to be more labile than DOC derived from terrestrial organic matter leaching, which is usually more recalcitrant and humified (Huguet et al., 2009; McKnight et al., 2001; Petrone et al., 2011). Consequently, the elevated humification index in the less urban watershed, RRN, with watershed-level SWM systems could have resulted from increased allochthonous inputs of recalcitrant terrestrial organic matter (Duan et al., 2014).

Differences in NO\textsubscript{3}\textsuperscript{−} sources among urban watersheds likely result from differences in the age of development and extent of % ISC and less likely due to restoration or management. High δ\textsuperscript{15}N-NO\textsubscript{3} isotope levels are indicative of nitrate from wastewater sources (Divers et al., 2014; Kaushal et al., 2011). NO\textsubscript{3}\textsuperscript{−} from wastewater was highest in one of the more urban sites (PMR), and because there are no point sources for wastewater in the streams of this study, this indicates greater NO\textsubscript{3}\textsuperscript{−} contributions from leaky sanitary sewers at this site (Divers et al., 2014; Kaushal et al., 2011); yet all sites showed wastewater as the greatest source of NO\textsubscript{3}. Due to stream restoration at MBR, the neighboring sewer pipes were repaired and stabilized (Dohey et al., 2006; Mayer et al., 2010; US EPA, 2009), likely resulting in less sewer leaks at Minebank Run longitudinally along the restored reach compared to the more urban streams (DRN and PMR). Nonethe-
less, during summer baseflow, the $\delta^{15}$N-NO$_3^-$ isotope levels were consistently high along each stream length, suggesting the influence of leaky sewer inputs through groundwater recharge (Divers et al., 2014; Hall et al., 2016); but during the rainier spring season, the more urban streams (DRN and PMR) showed a decline in $\delta^{15}$N-NO$_3^-$ isotope levels, indicating possible dilution of sewer-sourced nitrate from rainwater entering from connected impervious surfaces (Divers et al., 2014). This dilution of wastewater NO$_3^-$ was not observed at the other sites, potentially due to less connected impervious surfaces at the least urban watershed (RRN) and the reduction of peak discharge due to the reconnected floodplain for the restored stream (MBR) (Bohneke et al., 2002; Boyer and Kieser, 2012; Cendon et al., 2010; Hester and Goossen, 2010; Poff et al., 2006b). Nitrification was the second highest source for NO$_3^-$ at all sites, and likely contributed more NO$_3^-$ in the restored stream (MBR) and the least urban stream with watershed-level SWM systems (RRN) compared to the more urban streams (DRN and PMR), due to less labile carbon (Strauss and Lamberti, 2002) and possibly due to higher sediment C content at these sites (Arango and Tank, 2008). The greater atmospheric NO$_3^-$ during high flows in PMR and DRN is a result of the higher impervious surface cover at these sites, allowing for the more direct connection of rainfall to the stream corridor (Buda and DeWalle, 2009; Burns et al., 2009; Silva et al., 2002). Furthermore, the inverse relationship between $\delta^{15}$N-NO$_3^-$ and $\delta^{18}$O-NO$_3^-$ at all sites indicated mixing of sewage and atmospheric NO$_3^-$ to varying degrees among these urban watersheds (Kaushal et al., 2011). The downstream increase in $\delta^{18}$O-NO$_3^-$ after a spring rain event shows how the more urban streams maintain atmospheric NO$_3^-$ throughout their stream length. The restored stream only showed atmospheric source NO$_3^-$ in its headwaters (which is more developed) but not further downstream. The least urban watershed with SWM systems, RRN, showed minimal or no atmospheric NO$_3^-$ signal throughout its entire stream length, corresponding with it having no directly connected ISC. Conversely, during summer baseflow, there were no differences in the atmospheric NO$_3^-$ signal along the stream length for all four watersheds. These results suggest the dynamic potential of urban streams to transform nitrate along the broader urban watershed continuum based on gradients in land use and infrastructure (Kaushal et al., 2014a).

4.3 Variability in carbon and nutrient exports among urban watersheds

The higher C exports in the urban watersheds with greater % ISC compared to the restored stream and the least urban watershed with SWM systems may be due to increased autochthonous C production (described above) and greater leaky sanitary sewers (Kaushal and Belt, 2012). Inputs of leaves and other organic materials from street trees and organic matter delivered by storm drains from impervious surfaces likely also contributed to higher C exports in the urban degraded watersheds (Kaushal and Belt, 2012). Differences may have also stemmed from altered in-stream processing and elevated gross primary production in more urbanized, degraded streams (Kaushal et al., 2014a). The restored stream also likely had lower C exports due to increased ability to retain and process carbon in transient storage zones, like pools, through hyporheic exchange, or in the reconnected floodplain (Bukaveckas, 2007; Groffman et al., 2005; Klocker et al., 2009; Mulholland et al., 1997; Pennino et al., 2014), whereas degraded urban streams that are highly eroded can have less transient storage areas to potentially store and process carbon (Kurth et al., 2015; Sudduth et al., 2011a). Previous work at nearby sites suggests that labile C export from urban watersheds has the potential to increase oxygen demand, alkalinity, and denitrification (Kaushal et al., 2014a; Newcomer et al., 2012). Relatively less work has quantified exports of organic C from urban watersheds (Bullock et al., 2011; Worrall et al., 2012). The C exports of the urban watersheds in the present study, ranging from 6 to 57 kg ha$^{-1}$ yr$^{-1}$, were within the range or higher than nearby forested watersheds in North America and elsewhere, which range from 10 to 100 kg ha$^{-1}$ yr$^{-1}$ (e.g., Aitkenhead-Peterson et al., 2005; Dillon and Molot, 1997; Hope et al., 1994; Mulholland and Kuenzler, 1979; Tate and Meyer, 1983).

The TN exports in this study, which ranged from 3 to 8 kg ha$^{-1}$ yr$^{-1}$, were generally equal to or higher than most urbanized watersheds, which range from 0.2 to 9 kg ha$^{-1}$ yr$^{-1}$ (Lewis and Grimm, 2007; Petrone, 2010; Sobota et al., 2009), though some urban watersheds can export TN as high as 30 kg ha$^{-1}$ yr$^{-1}$ (e.g., Line et al., 2002). TN exports in this study were also similar to the exports estimated in some of the same urban watersheds at the Baltimore LTER site during similar annual runoff (Kaushal et al., 2008a; Shields et al., 2008). Previous work has shown that annual runoff is a strong predictor of annual N exports in the Baltimore LTER watersheds (Kaushal et al., 2011; Kaushal et al., 2008a), and the relationship between runoff and N export rate varies significantly across a broad range of sites based on the degree of watershed urbanization (Kaushal et al., 2014b). The higher TN exports in the more urban sites (PMR and DRN) compared to the restored stream (MBR) may be due to various reasons, such as greater N inputs from leaky sewers in the more urban and older watersheds and/or greater N removal through denitrification in the restored stream due its hydrologically connected floodplains (Kaushal et al., 2008b; Klocker et al., 2009), alluvial wetlands, and greater hyporheic exchange (Bukaveckas, 2007; Harrison et al., 2011; Kaushal et al., 2008b; Roley et al., 2012). In fact, the stream restoration at MBR involved bank stabilization and some repairs of sewer pipes (Doheny et al., 2006; Mayer et al., 2010; US EPA, 2009) and consequently may have reduced sewer leaks, but detailed research is needed to evaluate the effects of sewer repairs on watershed N inputs. There were also higher peak flows and a greater proportion of nu-
Transient exports at higher flows, as indicated by the F75 metric for the more urban sites (PMR and DRN), similar to previous work (e.g., Horowitz, 2009; Shields et al., 2008). The lower proportion of N exports during higher flows for the restored stream (MBR) may be due to the connected floodplain attenuating higher flows (Bohnke et al., 2002; Cendon et al., 2010; Hester and Gooseff, 2010), as evidenced by the effective discharge results described above and due to less connected impervious cover (Poff et al., 2006b; Smith et al., 2013). The lower TN exports in the watershed with SWM systems (RRN) may be due to an extensive undeveloped riparian buffer (Mayer et al., 2007) and from its SWM systems in the watershed (Bettez and Groffman, 2012), which both can enhance N removal.

Relatively fewer studies of P exports in urban watersheds exist compared to those addressing N exports (Duan et al., 2012; Petrone, 2010). P exports from the present study, which ranged from 0.14 to 0.54 kg ha−1 yr−1, were similar to those reported elsewhere (e.g., Petrone, 2010). Though prior to improvements in wastewater treatment, urban watersheds impacted by sewage treatment plans were previously reported to export P ranging from 0.027 to 2.11 kg ha−1 yr−1 (Hill, 1981). Watershed P exports were also within the range reported by Duan et al. (2012) for Baltimore LTER watersheds, where the less urban, more managed watersheds typically showed lower TP and soluble reactive phosphorus exports. The higher exports of TP and PO4−3 at the more urban watersheds (PMR and DRN) may indicate greater inputs from leaky sewers and possibly from erosion of the stream channel (personal observations) due to flashier hydrology at these sites (Paul and Meyer, 2001). The lower TP exports in the restored stream may be due to increased hyporheic exchange and floodplain connection, which have been shown to increase P retention (Butturini and Sabater, 1999; Mulhoffland et al., 1997). Higher F− and I− concentrations and exports in the older, more urban, and less managed sites further suggest that there are water inputs from leaky drinking water pipes and sewers (Darcan et al., 2005; Gehr and Leduc, 1992; Xu et al., 2016). More work is necessary to track sources of P in urban watersheds.

4.4 Flashiness of water, carbon, and nutrient exports among urban watersheds

As expected, the streams with greater % ISC (PMR and DRN) showed more flashy hydrology and evidence that overland flow or storm drain inputs were a significant flowpath (as supported by the water and nitrate isotope mixing model results). In-stream restoration features of MBR may have contributed somewhat to dampening flood pulses by promoting floodplain reconnection (Bohnke et al., 2002; Cendon et al., 2010; Hester and Gooseff, 2010). However, the inconsistently lower hydrologic flashiness metrics for MBR compared to the more urban streams (PMR and DRN) may indicate stream restoration has variable hydrologic impact (e.g., Emerson et al., 2005; Sudduth et al., 2011b) depending on the storm size or specific features of the stormwater management. At RRN, the lower % ISC, higher % SWM, and larger watershed size likely contributed to reduced hydrologic flashiness by disconnecting impervious surfaces and promoting infiltration (Meierdierck et al., 2010; Baltimore County, Maryland Department of Environmental Protection and Sustainability).

The significantly more pulsed C and nutrient exports in the more urban watersheds (PMR and DRN) may be attributed to hydrologic variability and impervious surface cover. Dissolved C, N, P, F−, and I− exports in the more urban watersheds could have also been more variable due to runoff from impervious surfaces and/or increased contributions from storm drains (Bernhardt et al., 2008; Hatt et al., 2004) and elsewhere in the stream corridor (i.e., sewage leaks) during storms, as shown in other studies (Divers et al., 2014; Kaushal et al., 2011; Phillips and Chalmers, 2009). We also found pulses in atmospheric NO3− sources (as indicated by δ18O-NO3−) during storms in the more urban watersheds, similar to Kaushal et al. (2011). The attenuation of peak discharge due to stream restoration observed at MBR, which reconnected the stream with the floodplain is likely a large factor in why MBR had comparatively less pulses in C and nutrient exports. Also the stabilization and replacement of sewer pipes along the restored stream (Doheny et al., 2006; Mayer et al., 2010; US EPA, 2009) likely reduced the potential for C and nutrients to leak into the restored stream. Similarly, the upland stormwater management features and less % ISC at RRN likely helped to dampen the flows and pulses in C and nutrient exports at this site compared to the more urban sites (DRN and PMR).

Based on the nutrient duration curves and the F75 metrics, the more urban watersheds (PMR and DRN) had greater exports of N, P, and wastewater indicator anions (F−, I−) during higher flows compared to sites with lower % ISC and greater stormwater management (RRN) or stream restoration (MBR). Other studies also show elevated nutrient exports during higher flows in urban watersheds (Duan et al., 2012; Kaushal et al., 2014b; Shields et al., 2008). The higher C, N, and P exports during baseflow at the restored stream (MBR) and the least urban stream (RRN) compared to the urban degraded watersheds (DRN and PMR) likely correspond with there being less peak flow discharge, based on the hydrologic flashiness results, and also greater groundwater recharge at these sites, due to less impervious surface cover, greater SWM systems, or floodplain reconnection (Bohnke et al., 2002; Boyer and Kieser, 2012; Cendon et al., 2010; Hester and Gooseff, 2010). Overall, the sources, fluxes, and flowpaths of groundwater across streamflow should be considered in management efforts to improve stream restoration strategies for reducing nitrogen exports.
5 Conclusion

Our results demonstrate that stream restoration and sewers influence the local variability of C and nutrient sources and fluxes among urban watersheds within the same city. Urban sewer infrastructure also influences sources, fluxes, and flow paths of water, carbon, and nutrients over time and should explicitly be considered as part of the urban hydrologic cycle (Kaushal et al., 2014c, 2015; Risch et al., 2015). NO$_3^-$ isosites, C quality, and the fluoride and iodide tracer data suggest that sources of N and C within the stream corridor, such as leaky sanitary sewers and storm drain inputs, strongly influence the amount and timing of exports. Previous work has focused on upland stormwater management, but additional consideration of nonpoint sources in close proximity to streams and groundwater is also warranted in stream restoration strategies. Because gravity-fed sewers often follow stream channels, restored streams can be redesigned to better protect sewers from damage and further erosion as a management priority (Mayer et al., 2010). Consequently, effective management of urban streams may require upgrading or repairing leaks in sanitary infrastructure in the stream corridor to reduce these major sources, in combination with stream restoration or stormwater management strategies for dampening flashy hydrology and minimizing connected impervious surfaces in the watershed. These combined strategies could then help to reduce nutrient exports during both baseflow and stormflow.

Potential stream restoration strategies to reduce C and nutrient export include reducing the velocity of water and allowing overbank flow through floodplain reconnection, increasing retention of groundwater, providing sustainable sources of labile organic C, reducing imperviousness in the watershed, or daylighting streams. More research is needed to assess the effectiveness of stormwater retrofits in older urban watersheds at mitigating stream degradation and improving water quality. Managing nutrient export from aging urban watersheds will require better knowledge of sources across hydrologic variability, particularly due to groundwater contamination from leaky sewers and other urban piped infrastructure.

6 Data availability

Data used for the research in this paper is available through the 4TU.Centre (2016) at the following URL and DOI: http://dx.doi.org/10.4121/uuid:3636e6b7d-09dc-4a96-8d19-3ea6b9a7841 as well as from NOAA (2014) at http://www.ncdc.noaa.gov/cdo-web

Information about the Supplement

The following information can be found in the Supplement:
- additional details on methods;
- additional site information and site map;
- table of mean annual C and nutrient concentrations for each watershed;
- table of flashiness metrics for mean daily carbon, nitrogen, and phosphorus exports;
- table of flashiness metrics for routinely sampled concentrations;
- table of flashiness metrics for water and nitrate sources;
- table of flashiness metrics for carbon source metrics;
- flow duration curves for each site;
- comparison of nutrient concentrations over time at each site;
- water isotope comparison; and
- seasonal relationship between $\delta^{15}$N-NO$_3^-$ vs. $\delta^{18}$O-NO$_3^-$.

The Supplement related to this article is available online at doi:10.5194/hess-20-3419-2016-supplement.

Author contributions. This paper is based on work from Michael Pennino’s PhD dissertation. Michael Pennino collected water samples, conducted data analysis, and wrote the manuscript. Sujay Kaushal contributed to the study design, and provided helpful feedback on data analysis and manuscript writing. Paul Mayer contributed to study design, coordinated water sample laboratory analysis, and contributed to manuscript revisions. Ryan Utz helped with data processing and statistical analysis, and Curtis Cooper collected the data for the comparison of pre- and post-restoration hydrologic response and analyzed the relationship between effective precipitation (Ppt) and effective peak discharge ($Q_{pk}$).

Acknowledgements. This research was supported by EPA NNEMS Award 2010-309; the NSF Graduate Research Fellowship Program under grant no. DGE-1144243; NSF Awards DGE 0549469, EF-0709659, DBI 0640300, CBET 1058502, CBET 1058038, and EAR 1521224; NASA grant NASA NNX11AM28G; Maryland Sea Grant Awards SA7528085-U, R/WS-2, and NA05OAR4171042; and Baltimore Ecosystem Study LTER project (NSF DEB-1027188). We thank Tiana Pennino for field assistance, Garth Lindner for providing rating curves for two of the stream gauges, Stu Schwartz for extensive guidance on statistical analysis for...
nutrient export calculations, and Steve Stewart and Robert Hirsch from the Baltimore County Department of Environmental Protection and Sustainability for providing stormwater management data. Claire Welty and Andrew Miller provided help interpreting water isotope and streamflow data, respectively, and helpful suggestions and discussions on an earlier version of the manuscript. The US Environmental Protection Agency, through its Office of Research and Development, funded and managed, or partially funded and collaborated in, the research described herein. It has been subjected to the agency’s peer and administrative review and has been approved for external publication. Any opinions expressed in this paper are those of the author(s) and do not necessarily reflect the views of the agency, therefore, no official endorsement should be inferred. Any mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Edited by: A. Butturini
Reviewed by: two anonymous referees

References


Greenman-Pedersen Inc.: Minebank Run II stream restoration design report and 100-year floodplain impact analysis, Appendices, Laurel, MD, USA, 10 pp., 2003.


R Development Core Team: available at: http://www.R-project.org, last access: 23 September 2013.


