We would like to thank the reviewers and the associate editor for their time in providing detailed, constructive comments regarding this manuscript. We have combined our responses to both reviewers’ comments below, and believe that their contributions will lead to significant improvements. Both reviewers raised concerns about: 1) methodological details, 2) interpretation of results, and 3) terminology and clarity of ideas. As requested by the reviewers and the associate editor, we have responded to each of these comments below, and have incorporated all changes into the manuscript.

I. METHODOLOGICAL INFORMATION

Both reviewers have expressed concern about the level of detail provided in the methods section, and certain specific methodologies used. We have compiled their general comments and replied to their examples where more than a simple textual response was deemed necessary.

R1 Comment 1: This paper appears to lack some methodological information, some of which is important and makes it difficult to assess what you did. Some of these examples of this are listed below.

And

R2 Comment 3) Some parts of the methods need clarification (e.g. supersaturation, DOM sample preservation). In addition some parts of the methods seem unnecessary given the results that are presented

In the submitted version of this paper, we described the different terminology for gas saturation in stream water (saturation ratio, or xsCO₂, xsCH₄, and xsN₂O) on page 7, lines 19-24. We have added the following text to clarify this section:

“Super-saturation is defined as having a saturation ratio >1 or when xsCO₂, xsCH₄, or xsN₂O is >0.” (Page 10, line 12)

R2, P6, L16: 0.7micron-filtered samples stored for 2 weeks seems inappropriate for a DOM composition analysis. 0.2 micron filtering is usually preferred.

In the original manuscript, we describe DOM sample preservation and analysis on page 6, lines 16-21. Following filtration through pre-combusted 0.7μM glass fiber filters, samples were stored in amber glass vials at 4°C and analyzed within 2 weeks following collection.
To the authors’ knowledge, this is an appropriate and commonly utilized filtration procedure for DOM fluorescence metrics. Glass fiber filter pore sizes are not available below 0.7 μM, and smaller filter materials (such as 0.2 μM nylon) have the potential to leach out fluorescently active compounds and/or measurable amounts of dissolved organic carbon during filtration.

We have added the following text to clarify this method:

“Detailed methodology for optical properties and fluorescence indices can be found in Smith and Kaushal (2015), and numerous other studies have followed a similar filtration and storage procedure (Singh et al. 2014, Sing et al. 2015, Huguet et al 2009, Dubnick et al. 2010, Gabor et al. 2014). Fluorescently active DOM constitutes a wide range of lability. While some highly labile compounds may break down within hours of sample collection, more recalcitrant forms can remain stable for months. The ‘two week window’ is a convention meant to facilitate comparisons between sites, rather than a biologically based limit to storage (Personal communication, Rachel Gabor, Shuiwang Duan).”

R1: P4 L22-24 and Table 1: You may want to explain why you decided to treat these watersheds as four categories of two replicates each, rather than eight watersheds varying continuously along a few axes (impervious surface cover, development age, etc.). I think the reason of different discrete stormwater infrastructure design types going with developments built at different times makes sense; you just might want to state it a little more explicitly.

The reviewer’s understanding of our reasoning for treating watersheds as replicates of different categories is correct. We do attempt to explain the reasoning for development of infrastructure types (page 4, lines 22-24), and we have clarified this section as follows (Page 6, lines 16-21):

“We identified four categories based on distinct combinations of stormwater and sanitary infrastructure dominating the greater Baltimore region, based on maps of stormwater control structures, housing age, and intensive field scouting. We then selected eight first-order streams paired across the four categories. The first order stream sites each were located in half in Red Run and half in Dead Run, sub-watersheds of the Gwynns Falls (Fig. 1). We have abbreviated the categories based on the dominant infrastructure feature as follows: 1) stream burial, 2) inline stormwater management (SWM) wetlands, 3) riparian/floodplain preservation, and 4) septic systems (Table 1).”

We have additionally reviewed the remainder of the text to ensure that the infrastructure groupings are not described as a gradient.

R1: P4 L26-28: Over what time period (i.e. year(s), season(s)/month(s)?, times of day?) Actually, you should probably give much of this this information earlier than this section, and I don’t think you did.
We have added this information to the “Temporal Sampling of dissolved gases and stream chemistry” section (Page 7, lines 7-10): “Headwater stream sites were sampled every two weeks for both water chemistry and dissolved gas concentrations. Chemistry sampling took place for two years, between January 2013 and December 2014. Dissolved gas sampling took place between July 2013 and July 2014. Sites were visited between the hours of 9 AM and 2 PM.”

R1: P5 L5: How did you define a study reach? Approximately how long were study reaches? This information should come up in the previous section.

R2, P5, L5: Unclear what is meant by “study reach”. It has not been defined.

“Five dissolved gas samples were collected per stream on each date, along an established 20m study reach either upstream adjacent to the gaging station. Gas samples were collected at 0, 5, 10, 15, and 20m from the fixed starting point of the study reach.” (Page 7, lines 9-10)

R1: P5 L17 & P8 L21: “Estimated using Google Earth software” sounds a bit sketchy. If you must mention Google Earth, include a citation for the program. Ditto at 8(21)), and also, what’s the precision on the Google DEM, and why didn’t you use the lidar one mentioned in 5(19-20); is it not more precise?

This point was brought up by reviewers regarding 1) sampling locations, and 2) channel gradient (S) of our headwater stream sites, as well as the reaches where SF6 injection by Pennino et al. (2014) took place.

We did use Google Earth to identify the latitude and longitude of sampling locations, with reasonable confidence since sites were co-located with road crossings. We have cited Google Inc. here (Page 8, line 15).

In the original manuscript, we did estimate channel gradient of Pennino et al (2014)’s sites. In the revised version, we have re-calculated the channel gradient using a 1-meter resolution DEM derived from LiDAR surveys. This changed S for these sites, which in turn affected our estimate and uncertainty surrounding Cesc. Gas flux estimates have all increased in response to this change. Our description of these methods has changed as follows:

“We estimated S of headwater streams with GHG sampling sites by measuring the change in elevation along the stream above and below stream gaging stations. We determined the latitude and longitude of the stream gage, which was co-located with GHG sampling sites in Red Run and Dead Run using a Trimble GeoXH handheld 3.5G edition GPS unit (10cm accuracy). We then plotted this location atop a 1-m resolution LiDAR-based digital elevation model (DEM, Baltimore County Government, 2002) in ArcMap 10. Using low points in the DEM to represent the stream channel, we then selected one point above and one point below the stream gaging station and measured the distance between these two points along the stream channel with the ‘Measure’ tool. We then calculated S based on the change in elevation divided by distance. The slope
measurement reach overlapped with, but did not coincide exactly with the gas sampling reach in order to ensure measurable differences in elevation. We followed the same protocol to estimate \( S \) for reaches in Pennino et al (2014), except rather than estimating points above and below a gaging station, we determined the change in elevation over the specific reach where \( \text{SF}_6 \) injections took place. Latitude and longitude for the upstream injection point and distance downstream were provided by Pennino et al (2014) provided data on the latitude and longitude of their \( \text{SF}_6 \) injection reaches. “ (Page 12, line 23 - Page 13 line 7)

R2: P5 L17-20: There are multiple ways to make these calculations; what actual commands or tools did you use to do this?

We have clarified these methods as follows:

“ Sampling locations were designated pour points in the hydrology tools workflow. Because sampling points were always co-located with road crossings, we were able to acquire the latitude and longitude of sampling sites using Google Earth software (Google Inc. 2009). Watersheds were delineated using a 2-meter resolution DEM (Baltimore County Government, 2002). We first corrected the DEM for spurious depressions using the “Fill” tool in the ArcMap10.0 hydrology toolbox. Next, we calculated flow direction for each pixel of this filled DEM raster. We then used the Flow Accumulation tool to evaluate the number of pixels contributing to each downstream pixel. After ensuring that each pour point was co-located on the map streams (i.e. areas with flow accumulation >500 pixels), we used the ‘Watershed’ tool to delineate the pixels draining into each sampled location.” (Page 8, lines 13-20)

R1: P8 L14, 17, & 24) & P9 L1-4: What is \( K_{20} \)? You did not previously explain what \( GT \) (from \( K_{GT} \)) means in general terms, so if that explanation was supposed to translate; it does not do so effectively. Ditto with \( K_{SF_6} \) and plain \( K \); are those at ambient temperature?

We have clarified our description reaeration coefficient (\( K \)) for a given gas (\( G \)) and temperature (\( T \), as well as the gas transfer \( (k600) \) in the following lines, starting on page 11:

We calculated the gas flux rate using Eq. (5) where \( F_{GT} \) is the flux (g m\(^{-2}\) d\(^{-1}\)) of a given gas (\( G \)) at ambient temperature (\( T \)) and \( d \) is water depth (m). \( K_{GT} \) (day\(^{-1}\)) is the reaeration coefficient for a given \( G \) at ambient \( T \). Measured and equilibrium gas concentrations \([C_{str}]\) and \([C_{eq}]\) were calculated following equations 3 and 4, then converted to units of g m\(^{-3}\).

\[
F_{GT} = K_{GT} \times d \times ([C_{str}] - [C_{eq}])
\]  

(5)

We modeled \( K_{GT} \) for each site and sampling date using the energy dissipation model (Tsivoglou and Neal 1976). The energy dissipation model predicts \( K \) from the product of
water velocity \((V, \text{ m day}^{-1})\), water surface gradient \((S)\), and the escape coefficient, \(C_{\text{esc}}\), \((\text{m}^{-1}, \text{Eq. 6})\).

\[
K = C_{\text{esc}} * S * V
\]

\[(6)\]

\(C_{\text{esc}}\) is a parameter related to additional factors other than streambed slope and velocity that affect gas exchange, such as streambed roughness and the relative abundance of pools and riffles. The \(C_{\text{esc}}\) value used in this study was derived from 22 measurements of \(K\), made using the \(\text{SF}_6\) gas tracer method, carried out across a range of flow conditions in four streams within 5 km of our study sites and reported in Pennino et al. (2014). \(C_{\text{esc}}\) was calculated as the slope of the regression of \(K\) vs. \(S*V\) from data in Pennino et al (2014) and was assumed to be representative of our headwater stream sites in Dead Run and Red Run.

We calculated \(C_{\text{esc}}\) to be 0.653 m\(^{-1}\) \((n=22, r^2=0.42, p= 0.001)\). The 95% confidence interval of this \(C_{\text{esc}}\) based on measured \(K_{20,O2}\) values was ±0.359 m\(^{-1}\), which corresponds to ±55% of a given gas flux estimate. This estimate of \(C_{\text{esc}}\) from these nearby sites was assumed to be representative of the 8 stream reaches investigated in this study. Given the moderate range of uncertainty in \(C_{\text{esc}}\), as well as additional uncertainties associated with slope estimation and relating \(C_{\text{esc}}\) to different stream sites, gas flux estimates must be interpreted with caution.

Measurements of \(K\) were converted to \(K\) for each GHG (as well as \(O_2\) for general comparisons) by multiplying by the ratio of their Schmidt numbers (Stumm and Morgan 1981). \(K\) measured at ambient temperature was converted to \(K\) at 20°C \((K_{20})\) following Eq. 7.

\[
K_{20} = \frac{K_T}{11.0421^{T-20}}
\]

\[(7)\]

In order to compare re-aeration rates across sites and prior studies, we calculated the gas transfer velocity, \(k_{600}\), which is defined as \(K_{20,O2}\) multiplied by water depth, with units of m d\(^{-1}\). (Page 11 line 18- page 12 line 15)

**R1: P8 L20** You say you, “measure[ed] the change in elevation over a reach with a handheld GPS unit.” Isn’t elevation from GPS units usually rather unreliable? Describe the precision of your GPS unit.

We have clarified our description of how the stream channel slopes at our gaging stations were determined as follows: (Page 12, line 18- Page 13 line 2):

“We estimated \(S\) at each GHG sampling site by measuring the change in elevation over a reach. We determined the latitude and longitude of the stream gaging using a Trimble GeoXH handheld 3.5G edition GPS unit (10cm accuracy). We then plotted this location atop a 2m resolution LiDAR-based digital elevation model (DEM, Baltimore County Government, 2002) in ArcMap 10. Using low points in the DEM to represent the stream channel, we then selected one point above and one point below the stream gaging station and measured the distance between these two points with the ‘Measure’ tool. We then calculated \(S\) based on the change in elevation divided by distance.”
R1: P3 L28: Go ahead and be more specific than “water chemistry” if you can do so concisely.

We have changed this to “were sampled every two weeks for dissolved carbon and nitrogen concentrations as well as and dissolved gases.” (Page 7, lines 7-8)

R2, P5, L1: Please specify what blanks are here.
We have clarified that we collected three gas blanks by filling vials with 25mL of helium in the field (Page 7, line 17)

R2, P5, L26: Not sure this equation and the associated text are necessary according to the results shown later.
Reviewer 2 is correct that we do not discuss the results from this mass balance calculation later on and it could justifiably be removed. We have removed this, as well as panels (e) and (f) from Figure 5.

R2, P5, L29: What about minor tributaries? Define better what you mean by major tributary.
We have clarified in the text that we sampled tributaries contributing more than 5% of the discharge to the main channel at a given point along the stream network, however minor tributaries, contributing less than 5%, were not measured. (Page 8, line 5)

R2, P6, L10-12: Specify how TDN and DOC were analyzed.

We have clarified that ‘TDN’ was measured using the ‘TDN’ method, which consists of high temperature combustion in the presence of a platinum catalyst, and clarify that the ‘Shimadzu’ instrument was a “TOC Analyzer.” (Page 9, lines 7-10).

R2, P6, L29: Why use a new name for this index if BIX is the name normally used?
We have replaced ‘index of autochthonous inputs’ with ‘BIX’ throughout the manuscript.

R2 Comment 2) The role of external (non-in stream) and non-biological sources of GHG is not well considered in the manuscript. This may also make some calculations such as the index of aerobic and anaerobic respiration inaccurate.

We have added clarification of the potential non-in-stream sources of GHGs in the introduction (Page 3 line 2) and discussion (Page 21, line 3) of the revised manuscript.

R2, P7, L25 to P8, L11: This index seems controversial and needs clarifications. Not sure it can be really applied because apparently, it does not take into account external (non-in-stream) GHG sources and non-biological GHG sources.

Reviewer 2 is correct that AOU does not account for non-biological sources of GHGs. We will clarify this assumption about using the index on page 8, lines 10-11 where we define AOU. We have clarified that that AOU differentiates between aerobic CO₂ and all
other anaerobic or abiotic sources (and not anaerobic specifically). By using this index without an additional metric for abiotic CO$_2$, we must assume that the proportion of abiotic CO$_2$ is small and invariant across sites and dates sampled. Richey et al. (1988) justified the assumption that abiotic CO$_2$ was minimal in their systems as follows “At ambient conditions (pH 6-7, alkalinity of 500-1000 ueq), with dissolved free CO$_2$ of 100-150uM or higher, the CO$_2$ produced through respiration remains primarily as dissolved CO$_2$. Thus ionic equilibrium reactions can be neglected.” Richey et al (1998)’s justification is not valid in all cases for our study, as pH measurements varied widely from 4.81 to 8.9, and site-average CO$_2$ concentrations were lower than 100uM on 20 out of 152 sampling sites and dates, and alkalinity was not measured. CO$_2$ and pH were only both within this range on 36 out of 152 occasions. Among these observations, there remains a significant, positive linear relationship between xs CO$_2$ and xsN$_2$O (p= 8.36 x10$^{-15}$, r$^2$ = 0.83) across all sites. We have thus changed the terminology to “xsCO$_2$-AOU” rather than anaerobic CO$_2$ throughout the text.

We have clarified in the text (Page 11, lines 4-16) that this index does account for external (non-in-stream) CO$_2$ and O$_2$ sources, and this was our main reason for using the index. Regardless of whether CO$_2$ and O$_2$ are produced within the stream, in the soil, or along groundwater flowpaths, the ratio of these two gases within the stream will represent the relative abundance of CO$_2$ production to O$_2$ consumption along that flowpath. Richey et al. (1988) and Daniels et al. (2002) are two examples of freshwater-based studies that used this index to evaluate anaerobic CO$_2$ production in freshwaters.

R2, P11, L21: The term “anaerobic CO$_2$ concentration” seems erroneous. It does not make much sense. The same applies for anaerobic N$_2$O or CH$_4$ concentrations.

In the original draft of this manuscript, we defined ‘anaerobic CO$_2$ as xsCO$_2$-AOU*1.2. We now refer to this metric as (xsCO$_2$-AOU) throughout the manuscript instead of anaerobic CO$_2$ in order to acknowledge potential abiotic sources of CO$_2$. We would like to additionally clarify that AOU is not used for any other gases (CH$_4$ or N$_2$O) and we do not make mention to ‘anaerobic N$_2$O’ or ‘anaerobic CH$_4$’ because, unlike CO$_2$, these gases are not produced and consumed in direct proportion to O$_2$.

R2, P7, L23-25: Unclear. Please explain better how Cesc was estimated from SF6 additions.

We have clarified the definition of Cesc in the methods as follows: (Page 11, line 21-Page 12, line 6)

“We modeled K$_{GT}$ for each site and sampling date using the energy dissipation model (Tsirovglou and Neal 1976). The energy dissipation model predicts K from the product of water velocity (V, m day$^{-1}$), water surface gradient (S), and the escape coefficient, C$_{esc}$, (m$^{-1}$, Eq. 6).

$$K = C_{esc} \times S \times V$$

(6)

C$_{esc}$ is a parameter related to additional factors other than streambed slope and velocity that affect gas exchange, such as streambed roughness and the relative abundance of
pools and riffles. The $C_{esc}$ value used in this study was derived from 22 measurements of $K$, made using the SF$_6$ gas tracer method, carried out across a range of flow conditions in four streams within 5 km of our study sites and reported in Pennino et al. (2014). $C_{esc}$ was calculated as the slope of the regression of $K$ vs. $S*V$ from data in Pennino et al (2014) and was assumed to be representative of our headwater stream sites in Dead Run and Red Run.”

R2: Table 1: I do not think so many decimals are necessary for most of these variables. Table 2: “0.000” = “<0.001” or “<0.0001”?

We have determined the significance limit for statistical tests in this study to be 0.008 using a bonferroni correction, in which the number of tests performed (6) is divided by the normal 95% significance limit (0.05) yielding 0.008. We have therefore changed the number of significant digits in all p-value reporting to match this limit.

R2, Table 4: If some variables were log-transformed (e.g. logDOC: NO3), this should be indicated in the methods section.

We have clarified that the log of DOC: NO3- was used for statistical comparisons (Page 14, line 16)

II. STATISTICAL ANALYSES

Comment 2: In your statistical methods (section 2.4, “Statistical Analyses,”) you execute a number of models (linear mixed effects, stepwise linear regression, etc., yielding all the results in Table 2 and 5) testing similar or related things. This may constitute a statistical multiple comparisons problem, i.e. increased chance of Type I error (https://xkcd.com/882). Consider either combining models (e.g. in a structural equations modeling framework or similar) or correcting for this risk of error. At the very least, try to combine your categorical and continuous variables for into a single model for each gas.

Reviewer 1 expressed concern about the statistical approach of using two modeling approaches to examine controls on each gas species citing that this approach seems redundant. The authors acknowledge that using two separate approaches for the purpose of predicting gas saturation values would increase the chance of Type I error; however, this was not the aim of our approach. The two models were used to examine first, whether or not there was consistent variation in gases across the categorical comparisons of watersheds, and secondly to examine whether or not gases could be predicted based on broader gradients in physical or chemical constituents that existed across all sampling dates and locations. In response to this reviewer’s comments we have additionally incorporated a bonferroni correction to the p-value by dividing the 0.05 significance threshold by the number of models (6 models total, three for each gas), so that only tests with $p<0.0083$ are considered significant. This does not change our results.
III. INTERPRETATION OF RESULTS

R1 Comment 3: Some interpretations of your results, most but not all minor, don’t entirely make sense, or seem incomplete. For example:

R1: 12(16-17): Are you sure the “influence” is actually “indirect” on “biogeochemical processes in streams,” or does the “indirect” part really only apply to GHGs? It seems like those things listed are directly related to biogeochemistry in general.

We have clarified here that, while watershed infrastructure was not a statistically significant predictor of GHG saturation in streams, the gradients in DOC: NO$_3^-$ that we found across all infrastructure types was strongly correlated with GHG saturation. We interpreted this to mean that infrastructure may directly influence DOC and NO$_3^-$ loading to streams, and that this C:N stoichiometry is likely to be an important controller of GHG abundance downstream. (Page 18, lines 10-16)

We have also added mention of GHGs produced along groundwater flowpaths, or entering streams directly from leaky sewer lines (Page 19, line 9)

R1: P12 L23: Plain “nitrogen” or “inorganic nitrogen?”

We have changed the wording here to ‘inorganic nitrogen’ (P18, line 20)

R1: P13 L9-10: “stoichiometric conditions more favorable for denitrification” would be a DOC: nitrate ratio closer to 1:1? If that ratio is different in incoming groundwater, wouldn’t the N$_2$O:CO$_2$ ratio from that groundwater be correspondingly different as well?

We are not sure we follow the reviewer’s question here, however we see the need here to clarify our interpretation of DOC: NO$_3^-$ and CO$_2$:N$_2$O ratios.

DOC: NO$_3^-$ stoichiometry is one way to examine whether biogeochemical conditions are favorable for one microbial process over another, as Taylor and Townsend (2010) describe in their in-depth metadata analysis of DOC: NO$_3^-$ stoichiometry across a wide range of ecosystems. Helton et al. (2015) also provide a comprehensive review of the ways in which stoichiometry between inorganic N and organic C can be interpreted in various ecosystems. The implications of this stoichiometry at small spatial scales, such as the stream-groundwater interface of headwater streams, can be more complicated, however, and we agree with the reviewer that our interpretation could be explained more clearly.

As noted by Taylor and Townsend (2010), a DOC: NO$_3^-$ ratio of 1:1 is ideal for denitrification, while DOC: NO$_3^-$ much below 1:1 signifies conditions favorable for nitrification. While this ratio reflects the biogeochemical condition at the location/time the sample was collected, it is the result of processes occurring along the upstream
flowpath. In predominantly groundwater-fed streams, for instance, heterotrophic denitrification may consume significant proportion of DOC along groundwater flowpaths of a septic plume, thus drawing down the DOC: NO$_3^-$ of upwelling groundwater. Denitrification converts DOC to CO$_2$ and NO$_3^-$ to N$_2$ and N$_2$O. Numerous studies have shown septic plumes to have high concentrations of NO3- (e.g. Aravena et al. 1993). DOC concentrations are variable, but tend attenuate with depth in the aquifer and/or flow distance along the plume (Aravena and Robertson 1998; Pabich et al. 2001). For instance, Pabich et al. found high concentrations of DOC (>20 mg/L) in the upper part of a septic plume, with an exponential pattern of attenuation with depth. Consistently high NO$_3^-$- paired with attenuating DOC can result in a very low DOC: NO$_3^-$ ratio by the time groundwater reaches stream. These conditions at the stream-scale are more favorable for nitrification. Since nitrification is a chemoautotrophic process, consuming CO$_2$ while producing N$_2$O, we would expect to see a negative relationship, or no relationship between CO$_2$ and N$_2$O if nitrification were the dominant N$_2$O production pathway in a given watershed. Instead, we find positive correlations between CO$_2$ and N$_2$O in nearly all watershed sites (Figure 4a). We suggest therefore that denitrification may be producing N$_2$O in the groundwater in our septic-dominated sites, and drawing down DOC: NO$_3^-$ along groundwater flowpaths. This interpretation remains hypothetical, however due to a number of biotic and abiotic processes occurring at the same time. Further work measuring solutes and gases along a groundwater flowpath is necessary to identify the mechanisms producing high concentrations of N$_2$O.

R1: P13 L24-25: You’ve made a big jump here, from relatively high emissions in certain places to “globally significant.” Consider reminding your reader (“reminding” insofar as this should go in the introduction first; currently it’s all just missing) what it would take for these locally high emissions to be globally significant- what’s the relative global contribution of streams in general; how much of global streams is urban stream, etc. It might make more sense to think of the impacts of NO$_2$ emissions in the city in terms of local air pollution than global GHGs. You might also think about if your findings suggest anything new for general biogeochemistry, as opposed to just the GHG emission application.

Rather than focusing on global emissions, have re-framed our results to point out here, and in other parts of the manuscript, that diffuse emissions from urban streams constitute a previously unaccounted for source of N$_2$O and CH$_4$. It is currently unknown how significant this source is, although one study shows that, for N$_2$O, sanitary sewers could emit as much N$_2$O per capita as current estimates for secondary WWTP plants (Short et al. 2014). There is evidence that most of the N$_2$O-N found in these streams originates as wastewater, and our study adds insight into the magnitude and variability of biogenic gases in streams draining septic and sewer infrastructure.

We have also emphasized the point that greenhouse gas emissions from urban streams may represent an important export pathway, for C and N from stream networks. Our results suggest that gaseous losses may need to be considered in urban watersheds from the perspective of mass transport and watershed C and N budgets, citing Gardner et al. (2015). (Page 21, line 5)
Since we did not measure NO$_2$ emissions, we are unable to comment on whether or not streams are a source of that gas in our study sites.

R1: Missing: How did you analyze “longitudinal variability,” or the effect of “distance from watershed outlet,” on any of the response variables, i.e., the output of the method described in section 2.1.3? You make claims about the results of this survey in section 3.6 and display graphs derived from the data in Fig. 5, and then about the significance of these findings in 14(10-18). However, it’s never apparent that you did more than eyeball that data to assess spatial trends. Moreover, my eyeballing does not match your eyeballing; I don’t see Fig. 5 as reflecting the patterns you describe in the text.

R1: P14 L10-18: See comment after “missing,” in Comment 1; it is unclear if you did a statistical analysis to support these claims.

*We have added statistical analysis of longitudinal surveys as follows: (Page 15, lines 2-8)*

“We analyzed longitudinal data using multiple linear regressions in order to evaluate whether patterns observed in headwater sites were representative of the broader stream network. We compiled data from four surveys – Red Run and Dead Run in spring and fall – and used a stepwise linear regression approach to determine the significant drivers for each gas (Table 6). Covariates included log of drainage area above each point, watershed (Red Run vs. Dead Run), season (spring vs. fall), DOC concentration, DIC concentration, TDN concentration, log of discharge, location (tributary vs. main stem), DOC: TDN molar ratio, a TDN by Drainage are interaction term, and a DOC by drainage are interaction term. We used the stepAIC() function in R to determine the optimal model formulation, selecting the model with minimum AIC.”

R1: P15 L17-18: “Variation in nonpoint sources and flowpaths” is not really an independent variable you tested; you don’t know what in the watershed, but outside the stream, is driving anything, beyond a bit of inference about groundwater.

*We have removed this sentence in the conclusion.*

R1: Section 3.5 and Fig. 4b: Why do you think the slope directions of the lines in Fig. 4b so variable? Address this in discussion.

*We have added the following discussion of this result: (Page 20, lines 3-6)*

“Overall, the relationships between CH$_4$ and CO$_2$ were much weaker and more variable than the relationships between CO$_2$ and N$_2$O (Figure 4). While CO$_2$ and CH$_4$ are sometimes correlated in wetlands and rivers with low oxygen (Richey et al. 1998), this was not the case for our study sites. Instead, CO$_2$ and N$_2$O were highly coupled, suggesting prevalence of NO$_3^-$ as a terminal electron acceptor over CO$_2$.”
**R1: Table 5: You never interpret your K20 results in the discussion.**

*We have added description and discussion of our modeled K_{600} values in the results (Page 17, lines 15-22) as follows:*

“GHG emission rates were sensitive to differences in modeled k_{600}. Despite having medium to low gas saturation ratios compared with other sites, DRKV had the highest GHG emission rates on all dates. This is due in part to having the highest slope (0.10 m/m), and thus the highest modeled k_{600} (m day^{-1}). Our 37 estimates of k_{600} ranged from 2.4 to 122.6.1 m d^{-1}. Site-averages for k_{600} varied from 5.39± 0.73 to 28.0± 7.0 m day^{-1}. The median value for all k_{600} estimates was 13.24 m day^{-1}. This range of values and site-averaged values extends beyond that measured by Pennino et al. (2014) of 0.5 to 9.0 m d^{-1}. The discrepancy between Pennino et al. (2014)’s k_{600} measurements is driven by differences in channel gradient. Gradients in the present study ranged from 0.01 to 0.1, while Pennino’s ranged from 0.001 to 0.016 m d^{-1}. Channel gradient (S) is also the parameter with the greatest uncertainty, thus warranting cautious interpretation of our gas emission estimates.

*We have added discussion of the modeled k_{600} values as they relate to gas fluxes as follows: (Page 20 line 24)*

“While our measured N_{2}O saturation ratios were highly correlated solute concentrations and redox conditions (Table 4), emission rates sensitive to the gas transfer velocity (k_{600}), which varied by two orders of magnitude in our study (Table 6). “

**R1: P12 L26: Can you not distinguish (or at least venture an educated guess) between “C and N inputs and/or microbial metabolism,” based on measurements/calculation of these gases individually, together with those of other gases?**

We have added the following text regarding this point in the discussion: (Page 19, lines 3-9)

“We speculate that the location of infrastructure on the landscape may affect the relative importance of direct anthropogenic loading vs. microbial processes on DOC: NO_3^− ratios of stream water. For instance we found high concentrations of and NO_3^− and low DOC in streams draining septic systems. Much of this excess NO_3^− is likely from septic plumes, but the lack of DOC may be the result of microbial C mineralization along subsurface flowpaths. On the other end of the spectrum, very low NO_3^− and TDN in streams draining watersheds in the floodplain preservation category, which were also newly developed. In this case, the higher C:N may have been driven by lower N leakage rates as
well as improved ecological function of the preserved floodplain wetlands to remove any N that does enter the groundwater from stormwater or sewage leaks.”

R2: P2, L4: Land use can alter GHG emissions from streams not only through changes in drivers of stream metabolism. Changes in external GHG sources (e.g. groundwater inputs, soil leaching, point sources) and some geochemical reactions may also be important. In general, only part of GHG emissions from streams come from in-stream metabolism. This relevant aspect is not made sufficiently clear in this manuscript.

_We agree with Reviewer 2 that land use can alter external GHG sources to the stream, along with changing in-stream metabolism. In the present form of the paper, we make mention of external GHG sources on several occasions (i.e. ‘Section 1.2. Role of Sanitary Infrastructure section), however we do not specifically attempt to differentiate between external vs. in-stream GHG production as we do not have data to back up this type of analysis._

IV. DESCRIPTION OF STUDY DESIGN

R1 Comment 4: You refer several times to a gradient or continuum of stormwater infrastructure, but you never elucidate the relationships between or ordering of the infrastructure types that makes them constitute a gradient or continuum. Explain, up front and early. For example:

R1: P2 L29: Is the “along the urban watershed continuum” significant? Does something change along this gradient about the effect of the wetlands, or do you just mean “in urban watersheds?”

_We have removed this mention of the ‘urban watershed continuum’ in this case, as the paragraph was re-written._

R1: P1 L16: It is not immediately clear how these seemingly discrete categories constitute “a gradient of stormwater and sanitary infrastructure”- gradient along what axis, what variable?

_We have clarified here, and throughout the text that the watersheds were compared as infrastructure categories rather than a continuum. For example, (Page 1, lines 15-20)_

“We hypothesized that urban infrastructure significantly alters downstream water quality and contributes to variability in GHG saturation and emissions. We measured gas saturation and estimated emission rates in headwaters of two urban stream networks (Red Run and Dead Run) of the Baltimore Ecosystem Study Long-Term Ecological Research Project.. We identified four combinations of stormwater and sanitary infrastructure present in these watersheds, including: 1) stream burial, 2) inline stormwater wetlands, 3) riparian/ floodplain preservation, and 4) septic systems.”
R1: P3 L20-21: “Urban watershed continuum” again- is that just a way to refer to the stretch from the infrastructure in the headwater downstream a bit, or are the different kinds of infrastructure arranged along a continuum, or what?

The reviewer is correct that it is a way to refer to the flowpath from the infrastructure to the headwater downstream a bit. We have removed this instance of using the term in order to clarify meaning. We intend to describe the ways in which infrastructure is part of the stream network in urban watersheds – and how the infrastructure/stream interface may play a significant biogeochemical role at a watershed scale in urban ecosystems. We have focused our study on understanding the role of urban infrastructure on greenhouse gas dynamics in urban waterways. A growing body of work has shown that nutrient and carbon loads to streams, as well as the biogeochemical processes within flowing waters is related not only to land cover (% impervious surface, urban density, etc) but also urban infrastructure. Connectivity between runoff-generating water sources (groundwater, overland flow, shallow subsurface flow) and urban infrastructure (sewer lines, stormwater conveyance pipes, drinking water pipes, constructed wetlands, etc.) is likely to influence not only the anthropogenic inputs of C and N to waterways but also the relative importance of biotic interactions on C and N removal along flowpaths. Kaushal and Belt (2012) describe a conceptual framework of how urban-impacted flowpaths may influence downstream export of nutrients as the ‘Urban Watershed Continuum.’

R1: P5 L22: This is the closest thing to an explanation you’ve made so far, and it still doesn’t really make sense.

We have removed this sentence and reference to the urban watershed continuum throughout the manuscript.

V. VAGUE WORDING CHOICES

R1 Comment 5: You could improve this paper by reducing vague and occasionally careless diction. Sometimes this problem makes your meaning somewhat unclear. For example:

Both reviewers had concerns about some of the vague and unclear phrasing in sections of this paper. We respond here to their general comments as well as the specific examples from their line-by-line comments. Overall, we have clarified the key ideas underlying this paper in the introduction, provide more concrete details to back up statements about the literature, and link our interpretation of results more clearly to the figures and tables provided. Specific examples can be found in the following responses to reviewers’ comments below.

R2: Title: I have the feeling that something is missing in the title. Maybe the word “of” before “urban”?

We changed title to ‘Influence of infrastructure on water quality and greenhouse gas
R1: P2 L3-4: Consider fleshing out “globally significant” with some actual numbers? Also, if you have space, it might not hurt to explain very briefly how this impact of rivers and streams on GHGs was determined. It is unclear here whether the figures you cite include urban streams or not, and why. In other words, could knowing about urban stream GHGs make these fluxes more or less “globally significant?” Without this piece of information, it is unclear if all of the potentially contributing factors to urban stream GHG emissions that you describe in the rest of the paragraph are already accounted for in the currently accepted stream GHG numbers and you’re just partitioning sources, or if you might revise the numbers on stream GHG fluxes as a result of this study.

We have re-structured much of the introduction and included more details to justify the measurement of GHGs from streams (Page 2, lines 11-18) as follows:

“Streams and rivers are dynamic networks that emit globally significant quantities of CO\textsubscript{2}, CH\textsubscript{4} and N\textsubscript{2}O to the atmosphere. CO\textsubscript{2} emissions via flowing waters are equivalent to half of the annual terrestrial carbon sink (1.2 Pg CO\textsubscript{2}-C yr\textsuperscript{-1}, Cole et al. 2007; Battin et al. 2009). Stanley et al. (2016) recently demonstrated that flowing waters are significant CH\textsubscript{4} sources as well, emitting approximately 28 Tg yr\textsuperscript{-1}, which is equivalent to between 10 and 35% of emissions from wetlands globally (Bridgham et al. 2013). Approximately 10% of global anthropogenic N\textsubscript{2}O emissions are emitted from river networks due to nitrogen contamination of surface and groundwater (UNEP 2013; Ciais et al. 2013). There is evidence that these N\textsubscript{2}O estimates, based on IPCC guidelines, might be too low, given growing evidence of high denitrification rates in small streams with high NO\textsubscript{3}\textsuperscript{-} loads (Beaulieu et al. 2011). “

We have additionally clarified that some of these studies do take into account N\textsubscript{2}O emissions from urban areas indirectly, by using population to estimate N inputs to watersheds (Page 2, lines 21-26):

“As urban land cover and populations continue to expand, it is critical to understand the impacts on waterways, including C and N loading and GHG emissions. While N\textsubscript{2}O emissions from both urban and agricultural sources are taken into account in models based on estimated watershed DIN loading (Nevison et al. 2000; Seitzinger et al. 1998), measurements validating these estimates or estimates of CO\textsubscript{2} and CH\textsubscript{4} in urban watersheds are rare.”

R2 P1, L17: Unclear what is meant by “watershed continuum”. I think it would be more correct to speak about river network. This study focuses on the river and not on the whole watershed. This should be clear throughout the manuscript.

&

R1 P3 (20-21): “Urban watershed continuum” again- is that just a way to refer to the stretch from the infrastructure in the headwater downstream a bit, or are the
different kinds of infrastructure arranged along a continuum, or what?

These two comments are related to the term ‘urban watershed continuum.’ As noted above, we have removed references to the urban watershed continuum in this paper.

R1: P1(27-29): Your concluding sentence is rather vague; for a start, “influenced” could mean almost anything. Could you be a bit more specific about what the “influence” was and what the “implications” are?

R2: P1, L29: This last sentence of the abstract does not seem appropriate. It refers to emissions, which are not the focus of the manuscript. I would rather include a more conclusive sentence here.

These two comments refer to the final sentence of the abstract. We agree with the reviewers that our study does not focus on emissions and will remove the last part of this sentence starting with ‘with significant implications…’

We have replaced this section with the following concluding sentence:

“Despite a decline in gas saturation from the headwaters, streams remained saturated with GHGs throughout the drainage network, however, suggesting that urban streams are continuous sources of CO2, CH4, and N2O.” (Page 2, lines 3-5)

R2: P3, L20-24: Yes, but how much do streams contribute to whole watershed GHG fluxes?

R1: To put your results in context a bit better, see Gallo et al. 2014 (“Physical and biological controls on trace gas fluxes in semi-arid urban ephemeral waterways” in Biogeochemistry 121(1) pp.189-207). They did related measurements in ephemeral streams in urbanized deserts, with similar results. For just nitrous oxide emissions from urban streams, there are several more relevant papers; try searching “nitrous oxide urban stream,” in Web of Science if you can. (No, I am not Gallo et al.)

While it is beyond the scope of this manuscript to robustly quantify emissions from streams in this region we acknowledge that more context is necessary here to justify the scalability of our results. We have added mention of the relative contribution of rivers to GHG emissions from terrestrial ecosystems (agricultural and otherwise) in the introduction (Page 2, lines 11-18). We have also incorporated a citation to Gallo et al. (2014), as this study is highly relevant to the growing understanding of greenhouse gas production in urban aquatic environments. (Page, 2, line 21)

R2, P15, L25-28: I suggest the authors try to include more results-based conclusions and implications at the end of the paper. It also seems confusing that the authors emphasize wastewater here, when the paper is about streams and GIs.

We have changed the conclusions to reflect the extent of our paper- the relationship
between GHG saturation and water quality metrics, as well as the relatively high GHG emission rates which are equivalent to intensively managed agricultural landscapes (Page 22, lines 15-25)

R1: 1 L22: “These variables” refers to the “drivers of GHG dynamics,” “infrastructure categories,” or both? If it’s the former, I guess this line just verifies that “nitrogen stoichiometry” etc. are in fact “drivers of GHG dynamics” in this context (as expected); if “these variables” are the “infrastructure categories,” then it’s a much more novel finding.

We have clarified which variables we are referring to in the text as follows:

“Multiple linear regressions including DOC: NO$_3^-$ and other variables (DO, TDN, and temperature) explained much of the statistical variation in nitrous oxide (N$_2$O, r$^2$= 0.78), carbon dioxide (CO$_2$, r$^2$=0.78), and methane (CH$_4$, r$^2$=0.50) saturation in stream water.” (P1, line 23).

R2: P1, L23: Not sure these r$^2$ values are helpful here. It is not clear which statistical test was used.

We have clarified that the r$^2$ values refer to the results from multiple linear regression models for each gas. (P1, line 23-24).

R2: P1, L26: Again, unclear use of r$^2$ value.
We have removed the r$^2$ value in this line (P1, line 26).

R1: P2 L16-17: When you talk about GI here, are you proposing that all GI will have the same effects, at least in terms of direction of change in GHGs, or might effects differ depending on GI type?

We have clarified the potential varying roles of different forms of GI on GHGs as follows (Page 4, lines 14-17)

“The form of GI (i.e. stormwater control wetland vs. riparian/floodplain preservation) may also influence GHGs due to 1) differences in water residence time and oxygen depletion in wetland vs. floodplain soils, and 2) differences in watershed-scale N removal capacity of the two different approaches. Newcomer Johnson et al. (2014) found that riparian/floodplain reconnection was more effective at reducing N export from streams, compared with stormwater wetlands in Baltimore.

R1: P2 L20: “Source of uncertainty” for what? Do you just mean “uncertain,” or do you mean that this role could change our understanding of global fluxes from rivers, or what?

We have removed this sentence from the manuscript.
R1: P3 L10: Specify *anaerobic* nitrification; this is unclear until 12(29). With plain “nitrification,” it at first seems like N₂O must be a typo for NO₂⁻. You also need a source here for the description of nitrification; I don’t think Taylor and Townsend 2010 suffices.

We are not sure we understand the reviewer’s comment here, as we do not use the term ‘anaerobic nitrification’ in this paper. As described on page 5, lines 9-10, nitrification is a chemoautotrophic process, which oxidizes NH₄⁺ to NO₃⁻. CO₂ is consumed during this process, and N₂O is also produced as an intermediate in the NO₃⁻ oxidation process. We have cited Taylor and Townsend (2010) because they provide an excellent framework for determining whether an environment is more favorable to nitrification over denitrification based on the ratio of NO₃⁻ to DOC. We have additionally added a more general reference about nitrification in aquatic systems to the text (Schlesinger 1997).

R1: P3 L18: “GHG emissions”- what about them? “Increased GHG emissions?”

We have changed the wording here to “Increased GHG emissions” (Page 6, line 4)

R1: P12 L27 & P15 L13: Provide a citation for “’hot spots’” if you’re going to put it in quotes, so we can verify which definition of “hot spot” you mean. Also, decide if you’re going to say, “’hot spot’” or just “hotspot;” be consistent.

Upon reflection on Reviewer 1’s suggestions about this analysis, we have removed the term(s) ‘hot spot’ in this section the paper and consistently separate it into two words where we do use the term. In place of using the term ‘hot spot’, we have changed the wording as follows:

“Understanding the spatial variability in N₂O concentrations, as well as the processes responsible for N₂O production and NO₃⁻ removal in watersheds is useful for informing watershed management.” (Page 19, line 10).

R1: P15 L23: “Role” or “influence?” Sometimes your point could be stronger if you provided concrete numbers to back up your assertions. For example: P2 L13: What does “substantially” mean? Can you provide numbers as to the relative contributions of nonpoint and point sources?

We have backed up the claim that ‘Sewage may also contribute substantially to N₂O emissions from urban streams’ with the following text addition: (Page 3, lines 20-26)

“Several studies have documented that wastewater leakage from municipal sewers often accounts for more than 50% of dissolved N in urban streams (Kaushal et al. 2011; Pennino et al. 2016; Divers et al. 2013). While sanitary sewer lines are known to leak dissolved N, N₂O losses are not accounted for in greenhouse gas budgets of large WWTPs that these pipes feed into. Short et al. (2014) measured intake lines from three municipal WWTPs and estimated that N₂O emissions from sewer lines alone on the same order of magnitude (1.7 g N₂O person yr⁻¹) as current IPCC estimates for per-capita
emissions from secondary WWTPs. Their study demonstrates the importance of constraining biogenic gas emissions from streams, which flow alongside and may receive gaseous inputs from aging sanitary sewer lines.”

R1: P3 L22-24: How is human population relevant? Also, please contextualize “fastest form of land use change;” that statement alone isn’t really enough to ascertain significance. Is the magnitude of the change (i.e. first derivative of land use rather than second derivative) large? Is urban land use large, relative to other uses? Or do you think urban watersheds contribute disproportionately much to GHGs for their size, and so are significant globally even if small?

We will remove this sentence and clarify as follows: (Page 6, lines 10-12)

“We have changed the methods here slightly, removing the need to identify specific high flow sampling dates. (Page 13 lines 3-7)

“Pennino et al’s (2014) measurements of V during gas injections ranged from 0.02 to 0.15 m s⁻¹. V measured at headwater gaging stations in our sites ranged from undetectable to 0.34 m s⁻¹. In order to avoid extrapolation, we limited our estimation of gas fluxes to sampling sites and dates with V in the range measured by Pennino et al. (2014). These conditions corresponded to 37 measurements total, spread unevenly across the four headwater sites with complete rating curves (DRAL, DRKV, RRRB, DRGG). K estimates were restricted to five dates at DRAL, 18 dates at DRKV, 11 dates at RRRB, and three dates at DRGG. “

R1: P2 L9: Again, on “implications,” try to be less vague if you can do so concisely. “Increase or decrease” or “change the magnitude of?” “Alter seasonality of?” Etc.

During our re-structuring of the introduction, we have removed this sentence and clarified whether we expect an increase or decrease in GHG emission from streams in other instances.

R1: P15 L1: By “typologies” you mean “types?”

In the methods section (page 4, lines 20-24), we originally described the combinations of sanitary and stormwater infrastructure in each pair of similar watersheds as ‘typologies.’ This has been changed to ‘categories’ throughout the manuscript.
VI. TRANSITIONS, DEFINING TERMS, ETC.

R1, Comment 6: Remember to maintain coherence and clarity of the paper through clear transitions, linking similar ideas, defining terms the first time you mention them, etc. For example:

R2 Comment 1) Some strange terms are used throughout the text that could be avoided (e.g. “watershed continuum”, anaerobic concentration)

We have taken these reviewers comments into consideration and have made changes to the wording, definition of terms and transitions of ideas throughout the manuscript. We have included examples of our response to these concerns below.

R2: Abstract: You don’t describe your “longitudinal” results here (the ones along stream length).

We have included the following description of our longitudinal results in the abstract: (Page 2 lines 1-7)

“Longitudinal surveys extending form headwaters to third order outlets of Red Run and Dead Run took place in spring and fall. Linear regressions of this data yielded significant negative relationships between each gas with increasing watershed size, as well as consistent relationships between solutes (TDN or DOC, and DOC: TDN ratio) and gas saturation. Despite a decline in gas saturation between the headwaters and stream outlet, streams remained saturated with GHGs throughout the drainage network, suggesting that urban streams are continuous sources of CO$_2$, CH$_4$, and N$_2$O. “

R1 P2 L21-23: How do these numbers/methods for calculating global fluxes that you cite here compare to the ones in 2(2-3)?

We have removed this sentence from the manuscript.

R1 P3 L29 - P4 L2: The final sentence in this paragraph seems out of place. Maybe shift it to the start of the next paragraph and end with, “, which facilitated site selection,” or something? If you don’t move the sentence, at least go ahead and explain why this information store matters. I mean, I can guess, but I shouldn’t have to do so, or to wait until you bring it up again later. Maybe just collapse the first two paragraphs into one?

We have removed the last sentence of this paragraph.

R1 P4 L5-6: Clarify timing. Everything was put in place in the 1950s-1970s, and the aging and cracking is now (or rather, when this study was conducted)? Also, “between” or “from?”

We have removed this paragraph in order to streamline our description of the study
design.

R1 P4 L13: Remind us which eight streams- “...the eight streams studied drained...?”

*We have clarified that we are referring to the headwater stream sampling sites, which are paired across eight infrastructure categories (Page 6, lines 15-21).*

R1 P4 L14-20: Some of this description of what types of infrastructure were built when might go better in the introduction. Or at least, you might want to introduce the concept of change in design through time in the introduction.

*We have rearranged this section and removed the description of stormwater and sanitary infrastructure designs to the introduction (Pages 2-4)*

R1 P4 L12-16: This sentence has a bit of a run-on feel; consider breaking down. Also, does “stormwater infrastructure... encompass older designs” and the newer GI ones? The way the sentence breaks doesn’t suggest so. You could say, “We define stormwater infrastructure broadly to encompass older designs such as stormwater drainage networks and newer forms of ‘green’ stormwater infrastructure (GI),” and then define each in a sentence (or so) each.

*We have removed this sentence in order to streamline the description of our study sites.*

R1: P5 L20: Unclear how GIS calculations in previous sentence are used; abrupt transition back to “these surveys” is hard to follow.

*We have clarified this in the text as follows: (Page 8 lines 12-20)*

“We calculated the watershed contributing area above each sampling point and flow length from each sampling point to the watershed outlet using Hydrology toolbox in ArcMap 10. Sampling locations were designated pour points in the hydrology tools workflow. Because sampling points were always co-located with road crossings, we were able to acquire the latitude and longitude of sampling sites using Google Earth software (Google Inc. 2009). Watersheds were delineated using a 2-m resolution DEM (Baltimore County Government, 2002). We first corrected the DEM for spurious depressions using the “Fill” tool in the ArcMap10.0 hydrology toolbox. Next, we calculated flow direction for each pixel of this filled DEM raster. We then used the Flow Accumulation tool to evaluate the number of pixels contributing to each downstream pixel. After ensuring that each pour point was co-located on the map streams (i.e. areas with flow accumulation > 500 pixels), we used the ‘Watershed’ tool to delineate the pixels draining into each sampled location.”

R1 P5 L25: “Relative contributions of inflow” to groundwater?

*Upon reflection and based on the reviewer’s comments, we have decided to remove this hydrologic mass-balance analysis from the paper since it is not directly discussed or used*
to interpret GHG results.

R1 P12 L30 - P13 L1: Consider referencing figures here (and more elsewhere in the discussion) to make it easy for readers to look back at the ratios etc. that you mention.

We have added figures to this section of the discussion as follows:

“We found a strong positive relationship between N₂O saturation and CO₂ concentrations, suggesting that denitrification was the primary source of N₂O (Figure 5a). By contrast, very low DOC: NO₃⁻ ratios (Figure 2) in stream water with highest N₂O saturation (Figure 3a) suggest that nitrification was the dominant process at these sites.” (Page 19, line 13-15)

R1 P13 L27 - P14 L9 & P15 L5-9: Most of this information should go in the introduction. You can refer back to it here insofar as your findings update or add to it, but it’s unclear that they do. It does not seem entirely relevant here.

We have removed the paragraph about drivers of methane and moved background information about nitrogen loading and N₂O to the introduction. (Starting Page 4, line 18)

R1 P14 L31-32: You do not make it clear how this information about plants is relevant. Are you saying that some other type of plant within the waters you surveyed might be releasing methane in this way, but you didn’t measure it? There are no transitions into or out of this part about the plants, either.

We have removed this section on plants as it is not relevant to the paper.

R1 P15 (26-27): It is unclear how exactly this part about wastewater relates to your results. Either make your transitions more clear, or move this sentence to a different section.

We have included more detail about the role of wastewater in our study as follows (Page 22, lines 19-25)

“Our results suggest that N from septic plumes and sanitary sewer lines is the principal source of N₂O saturation in our study sties. Dissolved inorganic N is highly correlated with N₂O in our study sites, and the highest values are only present in watersheds with aging sanitary sewer infrastructure or septic systems. Our observations of N₂O saturation and emissions from urban and suburban headwater streams are comparable with streams and ditches in intensive agricultural watersheds (Harrison and Matson. 2003; Outram et al. 2012). These results suggest that streams draining medium to low-density suburban or exurban land cover are comparable to those in intensively managed agricultural areas in terms of N₂O emissions.”
**R1 P15(28):** You have not brought up the concept of mitigation before, and it isn’t immediately obvious if mitigation per se is the goal, or how your results translate to doing mitigation. Elaborate.

_We have removed the mention of mitigation in the conclusions, as this is not the focus of this study._

**R2 Comment 4) The dynamics of CO₂ are not considered in the discussion section**

_We have added discussion of CO₂ dynamics as they correspond with N₂O and CH₄ respectively. (Starting page 18, line 18)_

**R2 Comment 5) Reference to relevant recent studies on GHG dynamics in urban streams are missing (e.g. see Alshboul et al. 2016 Environmental Science & Technology 50: 5555-5563 DOI: 10.1021/acs.est.5b04923 and references therein)._ 

_We have cited this paper in the introduction (Page 3, line 19)_

**Technical corrections:** Again, numbers preceding comments refer to page number (line number). Please do not feel obligated to respond to all of these; just make sure you have them the way you want them in the final version.

1(30): “Infrastructure” misspelled. Also, consistent capitalization of keywords?

_We have fixed this spelling error and capitalized all key words consistently (Page 2, line 8)._ 

3(9-10): Instead of, “nitrification is a chemoautotrophic process that produces,” you could just say, “nitrification chemoautotrophically produces,” (and then switch “, and consumes” to “and consuming”) for brevity.

_We have changed the wording here as suggested by the reviewer (Page 5, line 14)_

3(27), 5(20), & 7(22-23): Is just sticking a web link in here appropriate? For 5(20) and 7(22-23) especially, I think you need proper citations.

_We have included citations for these websites in the text and will move the websites to the references section. We have maintained the website reference to beslter.org, as this is a permanent link and not a specific data source. (Page 10, line 24)_

3(28): “, which” would be more grammatically appropriate than “that.”

_We have changed ‘that’ to ‘which’ in this sentence (now Page 9 Line 9)_

4(7): “In-line?” Repeats throughout document- just make sure you want “in-line” and not “inline” or “in line.”
We have changed all instances of ‘in-line’ to ‘inline’ in the document.

4(11-12): Maybe “and” instead of “that are;” the phrasing of this sentence is a bit awkward. Also, I think you could avoid the passive tense of “are located” (“exist?”).

In our broader re-writing of the discussion, we have removed this wording.

4(26): “First-order streams” instead of “first order streams,” yes?

We have changed ‘first order’ to ‘first-order’ here (Page 1, line 15 & page 6 line 19) and throughout the manuscript.

4(27-28): I’m not sure why you repeat all the categories when you just said them and even said that you just said them. Also, here you capitalized the categories and put apostrophes around them, whereas you didn’t in the last sentence; pick a format, and be consistent.

We have re-written this section to clarify and consolidate descriptions of infrastructure categories (Page 6, lines 12-25)

4(32): “Septa” or “septum?”

We have changed the wording here to ‘septum’ (Page 7, line 15)

5(3), 7(16), 10(11), & 13(26): Remove tab for consistent paragraph formatting.

We have removed tabs throughout the manuscript for consistent formatting.

5(3-4): Consider rephrasing for clarity and brevity, e.g.: “A single stream water sample was collected in a 250 mL high-density polyethylene bottle at each site. One sample duplication rotated site each sampling date.”

We have rephrased this sentence as suggested by the reviewer (Page 7, line 20)

5(10): Unnecessary “to.”

We have removed this ‘to’

5(15-16): Can shorten slightly by removing passive tense, i.e. “USGS provided discharge data.” Also, consider providing a citation for the USGS data here.

We have changed the wording of this sentence as follows:

“A minimum of 10 points was measured along each cross section. Discharge data was
provided by USGS when samples were co-located with a gaging station.” (Page 8, line 10)

6(9): “To the University?”

We have added ‘the’ to this sentence (page 9, line 5)

6(12): “Underestimates” or underestimations? Also, what “it” refers to is a bit unclear.

We have changed ‘under-estimates’ to ‘underestimation’ and clarified that ‘it refers to the NPOC method. (P9, line 8)

6(13 & 24), 10(3), 11(23), & 13(7): “Via” and “vs.” need not be italicized.

We have removed italicization of these words throughout the manuscript.

6(16 & 19): Move “(DOM)” up to first use.

We have defined dissolved organic matter (DOM) in the methods section (page 9, line 4) and refer to ‘DOM’ for the remainder of the manuscript.

6(19-20 & 27-28): You essentially describe what molecular weight characterizes which source twice in a row, and do it better the second time; condense.

We have removed the first mention of molecular weight and condensed the explanation as follows (Page 9, line 24- page 10 line 4)

“The humification index (HIX) is defined as the ratio of emission intensity of the 435-480 nm region of the EEM to the emission intensity of the 300-345 nm region of the EEM at the excitation wavelength of 254 nm (Zsolnay et al. 1999; Ohno 2002). HIX varies from 0 to 1, with higher values signifying high-molecular weight DOM molecules characteristic of humic terrestrial sources. Lower HIX indicates DOM of bacterial or aquatic origin (Zsolnay et al. 1999). The autochthonous inputs index (BIX) is defined as the ratio of fluorescence intensity at the emission wavelength 380 nm to the intensity emitted at 430 nm at the excitation wavelength of 310 nm (Huguet et al. 2009). Lower BIX values (< 0.7) represent terrestrial sources, and higher BIX values (> 0.8) represent algal or bacterial sources (Huguet et al. 2009).”

7(4): “Eq.’s?” Maybe just write it out.

We have changed “Eq’s (2-4)” to “equations 3 and 4 here. (Page 10, line 6)

7(4): “Rations” or “ratios?” (Pretty sure you mean “ratios.”)

We have changed ‘rations’ to ‘ratios’ here (page 10, line 6)
7(5): If you must put a comma before “(µmol L⁻¹), I think you need one after too.

*We have removed the comma before µmol L⁻¹ (Page 9, line 25)*

7(11) & Table 1: Combine things in parentheses in “(Eq. 3) (Stumm and Morgan 1981).” Similar change needed at end of caption for Table 1.

We have done so here and in the caption of Table 1.

7(19): “-“ may be unnecessary.

We have removed the hyphen here.

8(4): “In” or “at?”

We have changed ‘in’ to ‘at here (page 11, line 9)

8(7): “From” or “by?”

We have changed ‘from’ to ‘by’ here. (Page 8, line 4)

8(8): “Were,” not “where.”

*We have changed ‘where’ to were here (page 12, line 14).*

8(8-9): “Would be indicative of” can be shortened to “would indicate” or even “indicates.” You could also remove, “other CO₂ sources, namely.”

*We have changed the wording here to ‘indicates’ (Page 10, line 1)*

8(26): “P= ” or “p=?”

We changed ‘p’ to lower case here. (P 12, line 8)

8(27): Provide units again for “±0.058.”

We have clarified that this value corresponds to ‘Cesc, units of m⁻¹ (P 12, line 8)

9(13): Escaped “).”

We have fixed the escaped ‘)’.

9(19-20): Lost sentence fragment.

We have removed this sentence fragment in our broader edits of this section.
11(9): Second comma unnecessary. Also, why “may be,” and only in second alternative explanation?

We have removed the second comma and chanted ‘may be’ to ‘is’ (Page 18, line 8)

12(16): “Typologies however,” should probably be, “typologies, however.”

We have changed the comma usage in this sentence and changed ‘typologies’ to ‘categories’ throughout the text.

12(22): You can shorten, “were present across all four infrastructure typologies (Fig. 4c), which suggests,” to “present across all four infrastructure typologies (Fig. 4c) suggest.”

We have changed the wording here to reflect the reviewer’s suggestion (P18, line 4)

12(30): “Concentrations suggest that” should be, “concentrations, suggesting that.”

We have made this change to the text (Page 18, line 20)

13(24): “Warrants,” not “warrant.”

We have changed ‘warrant’ to ‘warrants’ (Page 21, line 6)

14(23): “With DOC:NO3- while other” could use a comma in the middle (i.e. “with DOC:NO3-, while other.”

We have removed this sentence during broader edits of the discussion section.

15(1-2): Isn’t there just the one negative relationship? (“The negative relationship” instead of “negative relationships.”)

We have changed the wording here to ‘the negative relationship between CH4 saturation and TDN, suggest…” (Page 22, line 11)

Table 1: Header word spacing is awkward.

We have consolidated the wording of headers for better alignment.

Table 4: In caption, “* Indicate” should be something like, “A ‘*’ indicates,” based on comparable sentences elsewhere.

We have changed the wording of this figure caption as recommended.

Table 5: You may be missing some commas towards the end of the list in the
caption.

We have added commas to the caption as recommended.

**Figure 1:** “Sampling sites and black dots signify” should have a comma after “sites.”

We have added a comma here as recommended.

**Figure 2:** “Points signify data points,” in the caption is a bit confusing; consider removing the second “points.”

We have removed the second ‘points’ as recommended in this figure caption.

**Figure 3c:** I know it will mess with the clarity of your outliers, but consider some kind of log scale here; the differences between the actual boxes and whiskers are almost completely unapparent.

We have changed the CH$_4$ panel on this figure (now Figure 3b) to have a log-scale.

**Figure 3:** In caption, “box and whiskers signify the median, first and third quartiles,” is unclear phrasing. At minimum, I think “box” needs to be plural.

We have clarified the meaning of the box plots in this figure caption.

**Figure 5:** Consider combining identical keys for panels (e) and (f), and perhaps some of the identical axes across panels as well. Unpunctuated letters representing figure panels within the caption text, e.g. “in panels a through d signify a saturation,” are confusing; “a” is also a word. Also, more specific date here?

We have removed the panels related to hydrologic mass-balance in this figure, as this analysis was not used to describe patterns in GHGs.

**15(16):** “Of aquatic ecosystems” is in the middle of a list which relates to it (either end would make more sense), and the “as well as” and “significantly alter” seem unnecessary; commas would do.

We have changed the wording here as follows: “Variations in urban infrastructure (i.e. SWM wetlands, riparian connectivity, septic systems) influenced C:N stoichiometry and redox state of urban streams. These in-stream variables, along with potential direct sources from leaky sanitary sewer lines may contribute to increased GHG production and/or delivery to streams.” (Page 22, lines 17-19).

**15(25):** “Include” not “includes.”

We have changed this to ‘include’ (Page 20, line 25)
R2, P11, L1: This subtitle is repeated 3 times in this page.
We have fixed the subtitles to match the text of each section.

R2, P7, L13: Remove “and” before “flux”?
We were not able to find this error in the submitted manuscript.

R2, Table 4: If some variables were log-transformed (e.g. logDOC:NO3), this should be indicated in the methods section.
We have added mention of the log DOC: NO3 ratio (Page 14, line 16)

R2: For greater clarity, I suggest keeping the same order for the 3 solutes (CO2, CH4 and N2O) in all tables and figures as well as in the text.
We have changed the order of gases throughout the figures and text.
&
2(25-26): For clarity, consider something like, “In urban watersheds, these factors likely vary with stormwater and sanitary sewer...”

Both comments refer to the same sentence. We have removed this sentence from the manuscript during our re-write of the introduction section.

3(5): Consider ending this sentence with an “as well,” or similar to tie back to previous sentence.
We have added ‘as well to the end of this sentence (Page 11, line 16).

2(24-25): Consider “Some key differences between the watershed types that might affect this relationship include,” for clarity. Alternatively, “...may differ substantially between urban and agricultural watersheds due to contrasting biogeochemistry and hydrology. Some key differences...”

We have removed this sentence from the manuscript during our re-write of the introduction.

14(16): “Detailed information” is not in itself a “step;” you need a verb, e.g. “Finding detailed information.”

We have removed this sentence during the process of editing the discussion section.

15(9-11): This sentence goes with the end of the last paragraph.
We have removed this redundant sentence.

14(5): By “relative proportion of different gases,” do you actually mean “methane production?”
We have changed the wording here as follows:
“As with CO$_2$, and N$_2$O, CH$_4$ saturation was negatively correlated with DO, however CH$_4$ was positively correlated with DOC: NO$_3^-$. CO$_2$ and N$_2$O, by contrast, were more strongly and positively correlated with TDN (Table 4). These patterns suggest that, along with redox conditions, carbon availability may modulate CH$_4$ production as well. (Page 21, lines 21-26).


We have moved this sentence to the introduction, and changed the wording to ‘The form’ rather than ‘Varying forms’ (Page 4, line 14)

12(30): “The source,” or just “the primary source,” or “a source?”

We have moved this sentence to the introduction (Page 3, line 20) and changed our wording here to be ‘the primary source.’

15(1): By “variations” you mean “differences?”

We have changed ‘variations’ to ‘differences’ here. (Page 21, line 19)

15(6): “Methodology” or “assumptions” (or “methodological assumptions”)?

We have removed this sentence during other edits of the discussion.

15(20): “Ecological?” What does that mean here?

We have removed the mention of GHG emissions from urban ecosystems in the concluding paragraph in order to focus more on results-based conclusions (i.e. connections between water quality and GHGs)

4(8): “Reflects” what? I think you mean the timing of development. Maybe rephrase: “...developed in the 2000s with more infiltration-based designs...”

We have re-written this section and removed the term ‘reflects’ from our description of this infrastructure category.

4(18): Maybe “...exists in various forms, including gravity sewers and septic systems, as well as a gradient...” or “...exists as both gravity sewers and septic systems along a gradient...” rather than the current, more ambiguous, “...exists in varying forms (gravity sewers and septic systems) as well as a gradient....”

We have removed this paragraph in our restructuring of this section.

11(28): “Consistent along the drainage network for Red Run and Dead Run”: do you mean looking intra-Red Run drainage network and intra-Dead Run drainage
network, or are you looking at both together as part of a larger drainage network? I think you mean the former, but your phrasing is unclear?

We have edited this paragraph as follows (Page 21, line 15):

“Synoptic surveys of N$_2$O saturation in Red Run and Dead Run in this study provide evidence that the entire network is a net source of N$_2$O (Fig. 5). N$_2$O saturation shows a significant decline with increasing drainage area (Table 6, Fig. 5), suggesting that emissions outpace new sources to the water column.”
References Cited


Influence of infrastructure on water quality and greenhouse gas dynamics in urban streams

Rose M. Smith¹², Sujay S. Kaushal², Jake J. Beaulieu³, Michael J. Pennino⁴⁵, Claire Welty⁵

¹Department of Biology, University of Utah, Salt Lake City, UT, 84112, USA
²Department of Geology & Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, 20742, USA
³U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Cincinnati, OH, 45220, USA
⁴U.S. Environmental Protection Agency National Health and Environmental Effects Research Lab, Corvallis, OR, 97333, USA
⁵Department of Chemical, Biochemical, and Environmental Engineering and Center for Urban Environmental Research and Education, University of Maryland Baltimore County, Baltimore, MD, 21250, USA

Correspondence to: Rose M. Smith (rose.smith@utah.edu)

Abstract. Streams and rivers are significant sources of nitrous oxide (N₂O), carbon dioxide (CO₂), and methane (CH₄) globally, and watershed management can alter greenhouse gas (GHG) emissions from streams. We hypothesized that urban infrastructure significantly alters downstream water quality and contributes to variability in GHG saturation and emissions. We measured gas saturation and estimated emission rates in headwaters of two urban stream networks (Red Run and Dead Run) of the Baltimore Ecosystem Study Long-Term Ecological Research Project. We identified four combinations of stormwater and sanitary infrastructure present in these watersheds, including: 1) stream burial, 2) inline stormwater wetlands, 3) riparian/floodplain preservation, and 4) septic systems. We selected two first-order catchments in by each of these categories, and measured GHG concentrations, emissions, and dissolved inorganic and organic carbon and nutrient concentrations bi-weekly for one year. From a water quality perspective, the DOC: NO₃⁻ ratio of streamwater was significantly different across infrastructure categories. Multiple linear regressions including DOC: NO₃⁻ and other variables (DO, TDN, and temperature) explained much of the statistical variation in nitrous oxide (N₂O, r²=0.78), carbon dioxide (CO₂, r²=0.78), and methane (CH₄, r²=0.50) saturation in stream water. We measured N₂O saturation ratios, which were among the highest reported in the literature for streams, ranging from 1.1 - 47 across all sites and dates. N₂O saturation ratios were highest in streams draining watersheds with septic systems and strongly correlated with TDN. CO₂ saturation ratio was highly correlated with N₂O saturation ratio across all sites and dates, and CO₂ saturation ratio ranged from 1.1 to 73. CH₄ was always super-saturated with saturation ratios ranging from 3.0 to 2,157. Longitudinal surveys extending from headwaters to third order outlets of Red Run and Dead Run took place in spring and fall.
Linear regressions of this data yielded significant negative relationships between each gas with increasing watershed size, as well as consistent relationships between solutes (TDN or DOC, and DOC: TDN ratio) and gas saturation. Despite a decline in gas saturation between the headwaters and stream outlet, streams remained saturated with GHGs throughout the drainage network, suggesting that urban streams are continuous sources of CO\textsubscript{2}, CH\textsubscript{4}, and N\textsubscript{2}O. Our results suggest that infrastructure decisions can have significant effects on downstream water quality and greenhouse gases, and watershed management strategies may need to consider coupled impacts on urban water and air quality.

**Key Words:** Greenhouse Gases, Urban Streams, Infrastructure, DOC, Nitrate, Methane, Carbon Dioxide, Methane, Nitrous Oxide

1 Introduction

Streams and rivers are dynamic networks that emit globally significant quantities of CO\textsubscript{2}, CH\textsubscript{4}, and N\textsubscript{2}O to the atmosphere. CO\textsubscript{2} emissions via flowing waters are equivalent to half of the annual terrestrial carbon sink (1.2 Pg CO\textsubscript{2}-C yr\textsuperscript{-1}, Cole et al. 2007; Battin et al. 2009). Stanley et al. (2016) recently demonstrated that flowing waters are significant CH\textsubscript{4} sources as well, emitting approximately 28 Tg yr\textsuperscript{-1}, which is equivalent to between 10 and 35% of emissions from wetlands globally (Bridgham et al. 2013). Approximately 10% of global anthropogenic N\textsubscript{2}O emissions are emitted from river networks due to nitrogen contamination of surface and groundwater (UNEP 2013; Ciais et al. 2013). There is evidence that these N\textsubscript{2}O estimates, based on IPCC guidelines, might be too low, given growing evidence of high denitrification rates in small streams with high NO\textsubscript{3}\textsuperscript{-} loads (Beaulieu et al. 2011).

While much of the research on GHG emissions from streams has taken place in agricultural watersheds, urban-impacted river networks receive similar N loads and have also shown elevated GHG concentrations and emissions (e.g. Daniel et al. 2001; Beaulieu et al. 2010, Beaulieu et al. 2011; Kaushal et al. 2014; Gallo et al. 2014). As urban land cover and populations continue to expand, it is critical to understand the impacts on downstream waters, including C and N loading and GHG emissions. While N\textsubscript{2}O emissions from both urban and agricultural sources are taken into account in models based on estimated watershed DIN loading (Nevison et al. 2000; Seitzinger et al. 2000), measurements validating these estimates or estimates of CO\textsubscript{2} and CH\textsubscript{4} in urban watersheds are rare. Quantifying the variability, drivers, and sources of GHG emissions from streams will illuminate the biogeochemical processes and potential role of urban infrastructure on nutrient cycling, water quality, and GHG budgets.
1.2 Role of sanitary infrastructure

The form and age of stormwater and sanitary infrastructure within a watershed can influence stream water GHG emissions in several ways. GHGs may enter urban streams directly through buried stormwater and sanitary infrastructure, or form increased production within streams in response to nutrient loading and/or geomorphic changes. We investigated the role of infrastructure on GHG emissions from streams in order to evaluate these potential drivers of heterogeneity within urban watersheds. Sanitary infrastructure encompasses a wide array of systems to manage human waste. In developed countries, sanitary infrastructure includes a combination of septic systems, sanitary sewers, and sometimes, combined stormwater/sanitary sewers. Storm and sanitary sewer lines are present in areas with medium-to-high density development. Sanitary sewer or combined sewer network delivers waste to centralized wastewater treatment plants (WWTPs), which treat influent and release effluent into larger rivers or coastal zones. Sanitary, storm, and combined sewers tend to follow stream valleys (i.e. low points in the landscape), are often made of erodible materials such as terra cotta or concrete, and tend to crack or develop leaks. Leaks in sanitary sewer infrastructure can lead to chronic nutrient loading throughout stream networks (Divers et al. 2013, Kaushal et al. 2011; Pennino et al. 2016; Kaushal et al. 2015). Septic systems, primarily used in low-density residential areas, are designed to settle out waste solids and leach N-rich liquid waste into subsurface soils and groundwater. Sanitary sewer infrastructure may influence GHG abundance and emission from streams directly via diffusion of gases out of gravity sewer lines (Short et al. 2014), or indirectly by microbial processing along surface and subsurface flowpaths (Yu et al. 2013; Beaulieu et al. 2011). While the present study focuses mainly on first to third-order streams influenced by sanitary sewer lines or septic systems, it is also worth mentioning that WWTPs are known to be a source of CH₄ and N₂O in urban areas, and contribute point-source GHG loading to larger rivers and coastal areas (Beaulieu et al. 2010; Strokal and Kroeze 2014; Alshboul et al. 2016).

Sewage leaks are likely the primary source of N₂O emissions from small urban streams (Short et al. 2014). Several studies have documented that wastewater leakage from municipal sewers often accounts for more than 50% of dissolved N in urban streams (Kaushal et al. 2011; Pennino et al. 2016; Divers et al. 2013). While sanitary sewer lines are known to leak dissolved N, N₂O losses are not accounted for in greenhouse gas budgets of large WWTPs that these pipes feed into. Short et al. (2014) measured intake lines from three municipal WWTPs and estimated that N₂O emissions from sewer lines alone on the same order of magnitude (1.7g N₂O person yr⁻¹) as current IPCC estimates for per-capita emissions from secondary WWTPs. Their study demonstrates the importance of constraining biogenic gas emissions from streams, which flow alongside and may receive gaseous inputs from aging sanitary sewer lines.
1.3 Role of stormwater infrastructure

Stormwater infrastructure varies widely across and within cities. From stream burial in pipes to infiltration-based green infrastructure (GI) designs, stormwater management designs have evolved over time (Collins et al. 2010, Kaushal et al. 2014). In Baltimore, where this study took place, stormwater management installed prior to the 1970s consisted of concrete-lined channels and buried streams (Baltimore County Department of Planning, 2010). Areas developed during the 1990s and 2000s are characterized by a more GI-based design approach, including but not limited to upland detention ponds, infiltration basins, wetlands and bio-swales. Stream restoration projects and riparian zone protections have also been established, restricting development within 100m of the stream corridor for new developments (Baltimore Department of Planning, 2010).

The form of stormwater infrastructure – whether stream burial, infiltration wetland, or restored riparian zone – may contribute to GHG saturation of groundwater and streams. Stormwater-control wetlands and riparian/floodplain preservation may increase or decrease CH₄ and N₂O emissions from streams, depending upon how watershed C and N inputs are routed along hydro-biogeochemical flowpaths. For instance, if these forms of GI are successful at removing excess N inputs to streams, GI may reduce N₂O emissions from flowing waters. Alternatively, GI may increase both N₂O and CH₄ inputs to streams and thus emissions by facilitating anaerobic microbial metabolism (Søvik et al. 2006; VanderZaag et al. 2010). The form of GI (i.e. stormwater control wetland vs. riparian/floodplain preservation) may also influence GHGs due to 1) differences in water residence time and oxygen depletion in wetland vs. floodplain soils, and 2) differences in watershed-scale N removal capacity of the two different approaches.

1.4 Variables controlling GHG production in urban watersheds

Reach-scale studies in streams across biomes have demonstrated that GHG production and emission is sensitive to changes in nutrient stoichiometry, organic matter quality, redox state, and temperature (e.g. Bernot et al. 2010; Kaushal et al. 2014a; Beaulieu et al. 2009; Dinsmore et al. 2009; Baulch et al. 2011; Harrison and Matson 2003). Several studies have shown that infrastructure can influence solute loading and stoichiometry of streams, which could in turn increase GHG production. For instance, Newcomer et al. (2012) measured higher rates of N uptake and denitrification potential in streams with restored riparian zones compared with degraded, incised urban streams. In-stream N uptake is also consistently higher in daylighted streams compared with streams buried in pipes (Pennino et al. 2014; Beaulieu et al. 2015). Upland or inline stormwater wetlands and retention ponds provide additional locations for focused N removal in urban watersheds (Newcomer et al. 2014; Bettez et al,
Sanitary infrastructure (i.e. leaky sewer lines and septic systems) can also be a source of N via leaching into groundwater (Shields et al. 2008; Kaushal et al. 2015; Pennino et al. 2016). In previous studies, carbon quantity and/or organic matter quality was correlated with N uptake or removal in urban streams and wetlands (Newcomer et al. 2012; Pennino et al. 2014; Beaulieu et al. 2015; Bettez et al. 2012; Kaushal et al. 2014). Inverse relationships between dissolved organic carbon (DOC) and nitrate (NO$_3^-$) concentrations have been found to persist across a wide variety of ecosystems ranging from soils to streams to oceans (e.g., Aitkenhead-Peterson and McDowell 2000; Dodds et al. 2004; Kaushal and Lewis 2005; Taylor and Townsend 2010). Recently, inverse relationships between DOC and NO$_3^-$ have also been reported for urban environments ranging from groundwater to streams to river networks (Mayer et al. 2010; Kaushal and Belt 2012; Kaushal et al. 2014a). A suite of competing biotic processes may control this relationship, by either 1) assimilating or reducing NO$_3^-$ in the presence of bioavailable DOC, or 2) producing NO$_3^-$ regardless of DOC status (Hedin et al. 1998; Dodds et al. 2004; Kaushal and Lewis 2005; Taylor and Townsend 2010). The former category includes heterotrophic denitrification, which oxidizes organic carbon to CO$_2$ and reduces NO$_3^-$ to N$_2$O + N$_2$ (Knowles, 1982), and assimilation of inorganic N (Wymore et al. 2015; Caraco et al. 1998; Kaushal and Lewis 2005). In the second category, nitrification chemosynthetic produces NO$_3^-$ by oxidizing NH$_4^+$ and consuming CO$_2$. Nitrification also yields N$_2$O as an intermediate product, and has been shown to dominate N cycling processes in low-DOC environments (Schlesinger 1997; Taylor and Townsend, 2010; Helton et al. 2015). In urban watersheds, denitrification is often limited by DOC due to increased N loading and/or decreased connectivity with carbon-rich soils in the riparian zone (Mayer et al. 2010; Newcomer et al. 2012). C:N stoichiometry are likely to be affected by stormwater and sanitary sewer infrastructure designs as well (Søvik et al. 2006; Collins et al. 2010; Kaushal et al. 2011). Stormwater wetlands may promote anoxic conditions and increase C:N ratio of stream water by increasing flow through carbon-rich soils (e.g. Søvik et al. 2006; Newcomer et al. 2012). Stream burial can reduce C:N ratios, if streams are buried in storm drains (Pennino et al. 2016; Beaulieu et al. 2014). Leaky sanitary infrastructure may additionally reduce the C:N ratio, and/or alter the form of carbon in streams (Newcomer et al. 2012).

1.5 Study goals

The goal of the present study was to identify patterns and drivers related to GHG dynamics in urban headwater streams draining different forms of infrastructure (stream burial, septic systems, inline SWM wetlands and riparian/floodplain preservation). Although less considered compared with nutrient loading, increased GHG emissions may be an unintended consequence of...
urban water quality impairments and biogeochemical processes occurring within and downstream of urban infrastructure. A growing body of work has shown that nutrient and carbon loads to streams are related not only to land cover metrics (% impervious surface, urban density, etc.) but also urban infrastructure (Shields et al. 2008; Kaushal et al. 2014). Connectivity between runoff-generating water sources (groundwater, overland flow, shallow subsurface flow) and urban infrastructure (sanitary sewer lines, storm sewers, drinking water pipes, constructed wetlands, etc.) is likely to influence nutrient export and biogeochemical function of waterways. An improved understanding of the relationship between infrastructure type and biogeochemical functions is critical for minimizing unintended consequences of water quality management, especially as growing urban populations place greater burden on watershed infrastructure (Doyle et al. 2009; Foley et al. 2005; Strokal and Kroeze 2014).

2.1 Sampling Methods

2.1.1 Study Sites

This study took place in collaboration with the Baltimore Ecosystem Study Long-Term Ecological Research (LTER) project (www.besler.org). We identified four categories based on distinct combinations of stormwater and sanitary infrastructure dominating the greater Baltimore region, based on maps of stormwater control structures, housing age, and intensive field surveys. We then selected eight first-order streams paired across the four categories. The first-order stream sites each were located in half in Red Run and half in Dead Run, sub-watersheds of the Gwynns Falls (Fig. 1). We have abbreviated the categories based on the dominant infrastructure feature as follows: 1) stream burial, 2) inline stormwater management (SWM) wetlands, 3) riparian/floodplain preservation, and 4) septic systems (Table 1).

Sites in the ‘stream burial’ category (DRAL and DIRS) drain watersheds with streams contained in storm sewers. Sanitary infrastructure in these watersheds is composed of aged sanitary sewer lines, installed prior to 1970 (Baltimore County Department of Planning, 2010). Streams in the ‘inline stormwater management’ category (DRKV and DRGG) originate in stormwater ponds or wetlands and also flow adjacent to aging sanitary sewer lines. Streams in the ‘riparian/floodplain preservation’ category (RRRM, RRSM) drain watersheds with newer development (after 2000), upland infiltration wetlands, and 100 m wide undeveloped floodplains (Baltimore County Department of Planning, 2010). Sanitary sewers were constructed in these watersheds between 2000 and 2010 (Baltimore County Department of Planning, 2010). Sites in the ‘septic systems’...
2.1.2 Temporal Sampling of Dissolved Gases and Stream Chemistry

Headwater stream sites were sampled every two weeks for solutes (DOC, TDN, HIX, BIX) and dissolved gas (CO₂, CH₄, and N₂O) concentrations. Chemistry sampling took place for two years, between January 2013 and December 2014, and gas sampling took place between July 2013 and July 2014. Sites were visited between the hours of 9 AM and 2 PM. Five dissolved gas samples were collected per stream on each date, along an established 20 m study reach either upstream adjacent to the gaging station. Gas samples were collected at 0, 5, 10, 15, and 20 m from the fixed starting point of the study reach. Samples were collected by submerging a 140 mL syringe with a 3-way luer-lock and pulling 115 mL of stream water into the syringe. We added 25 mL of ultra-high purity helium, to the syringe in the field, then shook syringes vigorously for 5 minutes to promote equilibration of gases between aqueous and gas phases. After equilibration, 20 mL of the headspace was immediately transferred into a pre-evacuated glass vial capped with screw-top rubber septum (LabCo Limited, Lampeter, UK), then transported to the laboratory, where samples were stored at room temperature for up to four weeks prior to analyses. Water temperature and barometric pressure during the equilibration were recorded in the field. We collected three helium headspace blanks by injecting 25 mL of helium into pre-evacuated vials in the field.

We collected stream water samples in a 250 mL high-density polyethylene bottles, one sample per site. One sample duplicate sample was collected on each sampling date, and the site for duplicate sample collection rotated among sampling dates. Dissolved oxygen (DO) concentration and pH were measured at the upstream end of each study reach using a handheld YSI 550. A dissolved oxygen meter (YSI Inc. Yellow Springs, OH) and an Oakton handheld pH meter (Oakton Instruments, Vernon Hills, IL).

2.1.3 Longitudinal Sampling of Dissolved Gases

Longitudinal surveys were conducted in June 2012, March 2014, and December 2014 in Red Run and Dead Run. Longitudinal sampling started at the outlet of each major tributary (Dead Run or Red Run), and extended every 500 m upstream to include the four bi-weekly sampled headwater sites in each watershed (Fig. 1). During spring and fall months, solute and gas samples were collected along all major tributaries (>5% main stem flow) as well as every 500 m along the main stem of Dead Run and Red Run. Minor tributaries (<5% of main stem flow) were not sampled. Stream discharge was measured at each sampling point.
using a Marsh-McBirney Flo-Mate hand held velocity meter (Marsh McBirney Inc., Frederick, MD, USA). We used cross-sectional measurements of stream velocity and water depth to calculate instantaneous discharge at each sampling site. We measured velocity and depth at a minimum of 10 points at each cross section in order to properly characterize flow across the channel. Discharge data was provided by USGS when sampling sites were co-located with a USGS gaging station (U.S. Geological Survey 2014). We used data from these surveys to determine whether or not headwater streams have more variable GHG saturation than the higher order parts of the stream network. Reach-scale hydrologic mass balances were calculated along the main stem of Red Run and Dead Run from these synoptic surveys following methods detailed previously (Kausal et al. 2014a, Newcomer Johnson et al. 2014). Along each reach of the main stem, relative contributions of inflow were calculated following Eq. (1): 
\[ Q_{GW} = Q_{DS} + Q_{US} + Q_{TRIB} \]
where \( Q_{GW} \) is the net groundwater input, estimated by difference using field measurements of \( Q_{DS}, Q_{US}, \) and \( Q_{TRIB} \). QDS is discharge measured in the main stem (m³ s⁻¹) at the bottom of a reach.

2.2 Laboratory Methods

2.2.1 Dissolved Gas Concentrations

Samples of headspace equilibrated gas concentrations (CO₂, CH₄, and N₂O) were stored at room temperature for up to 1 month in air tight exetainer vials and transported to the EPA National Risk Management Research Laboratory, Cincinnati, Ohio for analysis. Concentrations of CO₂, CH₄, and N₂O were measured using a Bruker 450 (Billerica, MA, U.S.A) gas chromatograph equipped with a methanizer, flame ionization detector (FID), and electron capture detector (ECD). Instrument detection limits were 100 ppb for N₂O, 10 ppm for CO₂, and 0.1 ppm for CH₄.

2.2.2 Solute Concentrations

Water samples were transported on ice to the University of Maryland, College Park and filtered using pre-combusted 0.7 μm glass fiber filters within 24 hours. A Shimadzu TOC analyzer (Shimadzu Scientific, Kyoto Japan) was used to measure total dissolved nitrogen (TDN) and dissolved organic carbon (DOC). The non-purgeable organic carbon (NPOC) method was utilized for DOC, despite potential underestimation of volatile compounds because the NPOC method is insensitive to variations in DIC.
TDN was measured on the same instrument using the ‘TDN’ method, which consists of high temperature combustion in the presence of a platinum catalyst. Nitrate (NO$_3^-$) concentrations were measured via colorimetric reaction using a cadmium reduction column (Lachat method 10-107-04-I-A) on a Lachat flow injection analyzer (Hach, Loveland, CO).

2.2.3. DOM Characterization

Filtered water samples were analyzed for optical properties in order to characterize dissolved organic matter (DOM) sources. After filtering (0.7 µm GF/F), samples were stored in amber glass vials at 4°C for a maximum of two weeks prior to analyses. Detailed methodology for optical properties and fluorescence indices can be found in Smith and Kaushal (2015), and numerous other studies have followed a similar filtration and storage procedure (Singh et al. 2014, Singh et al. 2015, Huguet et al. 2009, Dubnick et al. 2010, Gabor et al. 2014). Fluorescently active DOM constitutes a wide range of lability. While some highly labile compounds may break down within hours of sample collection, more recalcitrant forms can remain stable for months. The two-week window is a convention meant to facilitate comparisons between sites, rather than a biologically based limit to storage (Personal communication, Dr. Rachel Gabor & Dr. Shuiwnag Duan). Briefly, fluorescence and absorbance properties of DOM were measured in order to evaluate the relative abundance of terrestrial and aquatic sources to the overall DOM pool.

A FluoroMax-4 Spectrofluorometer (Horiba Jobin Yvon, Edison NJ, USA) was used to measure the emission spectra of samples in response to a variety of excitation wavelengths. Excitation-emission matrices (EEMs) were used for characterizing indices of terrestrial vs. aquatic DOM sources. The humification index (HIX) is defined as the ratio of emission intensity of the 435-480 nm region of the EEM to the emission intensity of the 300-345 nm region of the EEM at the excitation wavelength of 254 nm (Zsolnay et al. 1999; Ohno 2002). HIX varies from 0 to 1, with higher values signifying high-molecular weight DOM molecules characteristic of humic terrestrial sources. Lower HIX indicates DOM of bacterial or aquatic origin (Zsolnay et al. 1999). The autochthonous inputs index (BIX) is defined as the ratio of fluorescence intensity at the emission wavelength 380 nm to the intensity emitted at 430 nm at the excitation wavelength of 310 nm (Huguet et al. 2009). Lower BIX values (< 0.7) represent terrestrial sources, and higher BIX values (> 0.8) represent algal or bacterial sources (Huguet et al. 2009).

2.3. Calculations

Dissolved gas concentrations were calculated using equations 2-4. First, we used Henry’s law to convert measured mixing ratios (ppmv) to the molar concentration of each gas in the headspace vial [C$_g$] (µmol L$^{-1}$) following Eq. 2.
\[ [C] = \frac{P V}{RT} \]  
where P is pressure (1 atm), V is the measured partial pressure of the gas of interest (ppmv), R is the universal gas constant 
(0.0821 L atm mol\(^{-1}\) K\(^{-1}\)), and T is the temperature of a water sample (Kelvin) during headspace equilibration. We used Henry’s law and a temperature-corrected Bunsen solubility coefficient to calculate \([C_{aq}]\), the concentration of residual gas remaining in water following headspace equilibration (Eq. 3, Stumm and Morgan 1981)

\[ [C_{aq}] = \frac{V_p B_{Bunsen}}{RT} \]  
where V is measured gas mixing ratio (ppmv), Bp is the barometric pressure (atm), and Bunsen is the solubility coefficient in the vessel (L L\(^{-1}\) atm\(^{-1}\)). Calculations of the Bunsen coefficient were based on Weiss (1974) for CO\(_2\), Weiss (1970) for N\(_2\)O, and Yamamoto et al., (1976) for CH\(_4\).

The final stream water concentration \([C_{str}]\) was then calculated using mass balance of these two pools, described in Eq. (4), where \(V_{aq}\) and \(V_g\) were the volumes of water and gas respectively in a water sample with helium headspace.

\[ [C_{str}] = \frac{[C_{aq}] V_{aq} + [C_{g}] V_{g}}{RT} \]  
Because gas solubility is temperature dependent, it was useful to display gas concentrations as the percent saturation, or the ratio of the measured dissolved gas concentration to the equilibrium concentration. To determine gas saturation, the equilibrium concentration \([C_{eq}]\) was calculated based on water temperature, atmospheric pressure, and an assumed value for the current atmospheric mixing ratios of each gas following Eq. (3). We obtained current ratios for CO\(_2\) from The Keeling Curve (Scripps Institution of Oceanography, 2013), and N\(_2\)O and CH\(_4\) from the NOAA Earth Systems Research Laboratory (NOAA ESRL 2013; Dlugokencky, accessed 2013). Saturation ratio is defined as a ratio \([C_{aq}] / [C_{eq}]\) and excess (i.e. \(x\)CO\(_2\)) is described as a mass difference \((|C_{aq}| - |C_{eq}|)\). Supersaturation is the condition when the saturation ratio is greater than 1, or gas excess (i.e. \(x\)CO\(_2\)) is greater than 0.

### 2.3.2 Apparent Oxygen Utilization

Apparent oxygen utilization is defined as the difference between the O\(_2\) concentrations (µM) at equilibrium with the atmosphere vs. ambient measured O\(_2\) concentrations in the stream. A positive value of AOU represents net oxygen consumption conditions along the soil-groundwater-stream flowpath, while negative AOU (µM) represents net O\(_2\) production within the stream. Because aerobic respiration and photosynthesis couples CO\(_2\) production and O\(_2\) consumption, we can assume that AOU is equivalent to the CO\(_2\) produced / consumed along the same flowpath (Richey et al. 1998). Under aerobic conditions, respiration of organic
matter consumes O\textsubscript{2} and produces CO\textsubscript{2} at approximately a 1:1 molar ratio (Schlesinger 1997). Therefore, 1 mole of AOU should result in 1 mol of xCO\textsubscript{2} (Measured - equilibrium CO\textsubscript{2} concentration). This ratio was then used, with an offset to 1:2:1 to account for differences in diffusion constants for the two gases (Stumm and Morgan 1981; Richey et al. 1988), to determine the proportion of CO\textsubscript{2} produced by aerobic respiration. When CO\textsubscript{2} concentrations are greater than AOU, the difference between measured CO\textsubscript{2} and AOU (xCO\textsubscript{2}-AOU) represents additional sources from either anaerobic respiration or abiotic sources. We split our analysis of CO\textsubscript{2} into these two categories (AOU and xCO\textsubscript{2}-AOU) in order to determine whether patterns in CO\textsubscript{2} saturation were solely represented aerobic respiration or other processes and sources as well.

### 2.3.3 Greenhouse gas emissions

We calculated the gas flux rate using Eq. (5) where \( F_{\text{G}} \) is the flux (g m\(^{-2}\) d\(^{-1}\)) of a given gas (G) at ambient temperature (T) and \( d \) is water depth (m). \( K_{\text{GT}} \) (day\(^{-1}\)) is the re-aeration coefficient for a given G at ambient T. Measured and equilibrium gas concentrations \([C_{\text{eq}}]\) and \([C_{\text{str}}]\) were calculated following equations 3 and 4, then converted to units of g m\(^{-2}\). \( K_{\text{GT}} = \frac{F_{\text{G}}}{d + \left(\frac{C_{\text{eq}} - C_{\text{str}}}{K_{\text{G}}}ight)} \). (5)

We modeled \( K_{\text{GT}} \) for each site and sampling date using the energy dissipation model (Tsivoglou and Neal 1976). The energy dissipation model predicts \( K \) from the product of water velocity (V, m day\(^{-1}\)), water surface gradient (S), and the escape coefficient, \( C_{\text{esc}} \) (m\(^{-1}\), Eq. (3)).

\[
K = C_{\text{esc}} \times S \times V
\]  

\( C_{\text{esc}} \) is a parameter related to additional factors other than streambed slope and velocity that affect gas exchange, such as streambed roughness and the relative abundance of pools and riffles. The \( C_{\text{esc}} \) value used in this study was derived from 22 measurements of \( K \), made using the SE\textsubscript{G} gas tracer method, carried out across a range of flow conditions in four streams within 5 km of our study sites and reported in Pennino et al. (2014). \( C_{\text{esc}} \) was calculated as the slope of the regression of \( K \) vs. \( S \times V \) from data in Pennino et al. (2014) and was assumed to be representative of our headwater stream sites in Dead Run and Red Run.

We calculated \( C_{\text{esc}} \) to be 0.653 m\(^{-1}\) (n=22, r\(^2\)=0.42, p= 0.001). The 95% confidence interval of this \( C_{\text{esc}} \) based on measured \( K_{\text{esc}} \) values was ±0.359 m\(^{-1}\), which corresponds to ±55% of a given gas flux estimate. This estimate of \( C_{\text{esc}} \) from these nearby sites was assumed to be representative of the 8 stream reaches investigated in this study. Given the moderate range of uncertainty in
2.4 Statistical Analyses

2.4.1 Role of infrastructure and seasonality
A linear mixed effects modeling approach was used to determine the significant drivers of each gas across streams in different headwater infrastructure categories. Due to uncertainties in the gas flux parameters, GHG saturation ratios were used rather than GHG emissions to compare spatial and temporal patterns across sites. Mixed effects modeling was carried out using R (R Core Team, 2014) and the nlme package (Pinheiro et al. 2012) following guidance outlined in Zurr et al. (2009). Separate mixed effects models were used to detect the role of infrastructure category and date on each response variable. Response variables included saturation ratios for each gas (CO$_2$, N$_2$O, and CH$_4$), solute concentrations (DOC, DIC, TDN, NO$_3$), and organic matter source indices (HIX, BIX). Fixed effects were ‘infrastructure category’ and ‘sampling date,’ as well as an interaction term for the two. The effect of a random intercept for ‘site’ was included in each model. The statistical assumptions of normality, and equal variances were validated by inspecting model residuals. When necessary, variances were weighted based on infrastructure category to remove heteroscedasticity in model residuals (Zuur et al. 2009). The assumption of temporal independence was examined by testing for temporal autocorrelation in each response variable. This test was performed using the function ‘corAR1()’, which is part of the package ‘nlme’ in R. The significance of random effects, weighting variances, and temporal autocorrelation was tested by comparing Akaike information criterion (AIC) scores for models with and without each of these attributes. Additionally, pairwise ANOVA tests were run to determine whether each additional level of model complexity significantly reduced the residual sum of squares. Final model selection was based on meeting model assumptions, minimizing the AIC value, and minimizing residual standard error. Pairwise comparisons among infrastructure categories were examined using the Tukey HSD post-hoc test (lsmeans package, Lenth, 2016) for each response variable where ‘infrastructure category’ had a significant effect. Where ‘infrastructure category’ did not have a significant effect on a response variable after incorporating ‘site’ as a random effect, a separate set of linear models was run with ‘site’ and ‘date’ as main effects rather than ‘infrastructure category’. The role of ‘site’ was evaluated in these cases to determine the degree to which site-specific factors overwhelmed the effect of infrastructure category.

2.4. Role of environmental variables on gas saturation

A stepwise linear regression approach was used to examine the role of multiple environmental variables on CO$_2$, N$_2$O, and CH$_4$ saturation across sites and dates. Predictor variables were selected via backward stepwise procedure, using the ‘Step’ function in R. This involves first running a model that includes all potential driving factors, then running sequential iterations of that model after removing one variable at a time until the simplest and most robust combination of predictors was achieved. Model fit at
each step was evaluated using the AIC score. Parameters that did not reduce AIC when comparing models were removed until the model had the best fit with the minimum number of factors. The initial list of potential drivers included temperature, DO, DOC, TDN, DIC, HIX, and the BIX. Prior to the stepwise regression, we calculated the variance inflation factor (VIF) for each response variable to test for multicollinearity. VIF > 3 was the cut off for assessing multicollinearity. All variables in this study were below the VIF > 3 threshold (Zuur et al. 2010).

Analysis of covariance (ANCOVA) was carried out to determine whether relationships among gases (CO₂ vs. N₂O, CO₂ vs. CH₄) and solutes (log of DOC/NO₃ ratio) varied systematically across infrastructure categories. ANCOVA involved comparing two generalized least squares models. The first linear model included an interaction term between one of the predictor variables (i.e. DOC or CO₂) and infrastructure category to predict the response variable (N₂O or CH₄). The second was a linear model with the same two independent variables but no interaction term. When infrastructure category had a significant influence on both the intercept (first model) and slope (second model) of a relationship, this refuted the null hypothesis that infrastructure category had no influence on a relationship.

Because we used three separate models to evaluate variations in three GHG concentrations (for across infrastructure categories, continuous variables, and ANCOVA), we used a Bonferroni correction for the 95% confidence level. We determined the new confidence level by dividing the 95% level (0.05) by the number of models used on all gases across headwater stream sites (6). This new p-value (0.008) was then used to determine significance rather than 0.05.

### 2.4.3 Longitudinal variability in gas saturation

We analyzed longitudinal data using multiple linear regressions in order to evaluate whether patterns observed in headwater sites were representative of the broader stream network. We compiled data from four surveys – Red Run and Dead Run in spring and fall – and used a stepwise linear regression approach to determine the significant drivers for each gas (Table 6). Covariates included log of drainage area above each point, watershed (Red Run vs. Dead Run), season (spring vs. fall), DOC concentration, DIC concentration, TDN concentration, log of discharge, location (tributary vs. main stem), DOC: TDN molar ratio, a TDN by Drainage area interaction term, and a DOC by drainage area interaction term. We used the stepAIC() function in R to determine the optimal model formulation, selecting the model with minimum AIC.
3 Results

3.1 Effect of infrastructure on water quality and DOC: NO$_3^-$ ratios

We detected significant differences among TDN, NO$_3^-$, and DOC: NO$_3^-$ ratios across infrastructure categories (Table 2). TDN concentrations ranged from 0.12 to 8.7 mg N L$^{-1}$ (Table 3). Pairwise comparisons yielded significantly higher TDN concentrations in sites in the typology of ‘septic systems’, compared with the ‘in-line SWM wetlands’ typology, and sites in the ‘riparian/floodplain preservation’ typology. Sites in the ‘stream burial’ typology fell within the mid-range of TDN concentrations and were not different from any other category. DOC concentrations varied widely from 0.19 to 16.89 mg L$^{-1}$, but were not significantly predicted by infrastructure typology (Table 2). DOC: NO$_3^-$ ratios varied over four orders of magnitude, from 0.02 to 112 (Fig. 2). Infrastructure typology was a significant predictor of DOC: NO$_3^-$, with the lowest ratios in sites with septic systems and highest in sites with riparian/floodplain preservation (Fig. 2). Pairwise comparisons showed no difference in DOC: NO$_3^-$ ratios between in the in-line SWM wetland and complete stream burial typologies, however (Fig. 2).

3.2 Effect of urban infrastructure on DOM quality

Measurements of HIX ranged from 0.30 to 0.90 while BIX ranged from 0.40 to 1.15 across all sites and sampling dates in headwater streams. Streams draining septic system infrastructure had significantly lower HIX values than any other infrastructure typology. BIX values showed no significant pattern across infrastructure typologies (Table 2).

3.3 Effect of urban infrastructure on gas concentrations

Mixed effects models did not detect significant influence of infrastructure typology alone on CO$_2$, CH$_4$, and N$_2$O saturation in streams. There was, however, a significant interaction effect between sampling date and infrastructure typology on the saturation ratios of all three gases (Table 2). This indicated that sampling date was important to GHG saturation for some infrastructure typologies, or that the effect of infrastructure is dependent upon sampling date. The second set of linear models, which used site rather than infrastructure category as a main effect, yielded significant differences across all sites for N$_2$O (Fig. 3). Similarly, for CO$_2$, there were significant differences in 25 out of 28 pairwise comparisons. Pairwise comparisons across sites for CH$_4$...
saturation were significant in 23 out of 28 cases. These patterns suggest that site-specific effects overwhelmed the role of infrastructure categories on GHG saturation.

3.4 Effect of environmental variables on gas concentrations

Stepwise model parameter selection yielded several variables that correlate with each GHG saturation ratio (Table 4). TDN was the strongest predictor of N$_2$O saturation, followed by DO. The final model for N$_2$O ($r^2=0.78$) also included temperature, HIX, BIX, %SWM, and DOC:NO$_3^-$. CO$_2$ saturation had a similar pattern of predictors and nearly identical model fit ($r^2=0.78$). DOC:NO$_3^-$ ratio was the strongest predictor of CH$_4$ saturation followed by DO and temperature. HIX, %IC, and %SWM were also related to CH$_4$ saturation, but TDN and BIX were not.

3.5 Covariance among GHG abundance and C: N Stoichiometry

AOU ranged from -180.9 to 293.9 across all sites and sampling dates, however AOU was only negative (net oxygen production along surface and subsurface flow paths) in 6% of samples, or 43 out of 691 measurements. N$_2$O was significantly but weakly correlated with AOU ($p<0.008$, $r^2=0.12$). and strongly correlated with $\Delta$CO$_2$-AOU ($p<0.008$, $r^2=0.87$). Log of CH$_4$ saturation ratio was very weakly correlated with AOU ($p<0.008$, $r^2=0.01$) as well as $\Delta$CO$_2$-AOU ($p<0.008$, $r^2=0.07$). The relationships between $\Delta$CO$_2$-AOU and both N$_2$O and CH$_4$ saturation ratios were also significantly different between categories (Fig. 4).

There was an overall negative relationship between DOC and NO$_3^-$ with a significant interaction with infrastructure category (Fig. 4; ANCOVA p-value < 0.005).

3.6 Longitudinal Patterns in GHG saturation

Spatial variability in GHG saturation was examined in order to evaluate whether concentrations measured in tributaries were consistent between headwaters and larger 3$^{rd}$ order watersheds of Red Run and Dead Run respectively (Fig 5). Multiple linear regressions yielded a set of distinct controlling factors on saturation of each gas. The optimal models for CO$_2$ and N$_2$O were similar and included the log of drainage area, TDN concentration, log of discharge, and TDN x discharge interaction term. The CO$_2$ model also included DOC:TDN molar ratio. The optimal model for CH$_4$ saturation was slightly different, and included log
3.7 Greenhouse gas emissions

GHG emission rates were sensitive to differences in modeled \( k_{500} \). Despite having medium to low gas saturation ratios compared with other sites, DRKV had the highest GHG emission rates on all dates. This is due in part to having the highest slope (0.10 m/m), and thus the highest modeled \( k_{500} \) (m day\(^{-1}\)). Our 37 estimates of \( k_{500} \) ranged from 2.4 to 122.6.1 m d\(^{-1}\). Site-averages for \( k_{500} \) varied from 5.39± 0.73 to 28.0± 7.0 m day\(^{-1}\). The median value for all \( k_{500} \) estimates was 13.24 m day\(^{-1}\). This range of values and site-averaged values extends beyond that measured by Pennino et al. (2014) of 0.5 to 9.0 m d\(^{-1}\). The discrepancy between Pennino et al. (2014)’s \( k_{500} \) measurements is driven by differences in channel gradient. Gradients in the present study ranged from 0.01 to 0.1, while Pennino’s ranged from 0.001 to 0.016 m d\(^{-1}\). Channel gradient (S) is also the parameter with the greatest uncertainty, thus warranting cautious interpretation of our gas emission estimates.

Site-average CO\(_2\) emissions ranged 6.4± 2.3 g C m\(^{-2}\) day\(^{-1}\) at DRAL (+ standard error) to 134 ± 30.2 at DRKV. Mean emission rates for DRGG and RRRB were 11.5 ± 6.1 and 10.3 ± 1.7 respectively. Site-average CH\(_4\) emissions ranged from 2.6 ± 1.1 at DRAL to 102.5± 75.6 mg C m\(^{-2}\) day\(^{-1}\) at DRKV. N\(_2\)O emissions ranged from 5.1± 0.8 at RRRB to 149 ±33.9 mg N m\(^{-2}\) d\(^{-1}\) at DRKV. The full range of values and standard errors for fluxes are listed in Table 5.

4 Discussion

4.1 Overview

This study showed strong relationships between urban water quality and GHG saturation across streams draining different forms of urban infrastructure. N\(_2\)O and CO\(_2\) saturation was correlated with nitrogen concentrations, but did not differ between infrastructure typologies. DOC: NO\(_3\)-did not differ among the four infrastructure categories, however (Table 2). While infrastructure
categories did not show a significant predictor of GHG saturation in streams, the gradients in DOC: NO$_3$ found across all categories was strongly correlated with GHG saturation. Stoichiometric variation may thus serve as a predictor of GHG saturation downstream where land cover and infrastructure does not. While direct GHG loading to streams from leaky sanitary and/or stormwater infrastructure may play a role, the strongest predictors of GHGs in this study were continuous environmental variables (i.e. TDN and DOC concentrations, DO, temperature), rather than categorical (infrastructure category). Relationships between anaerobic $\Delta$CO$_2$ – AOU and N$_2$O saturation further suggest that anaerobic metabolism contributes to N$_2$O production along hydrologic flowpaths (Fig. 4).

4.2 C:N Stoichiometry as an Indicator of Microbial Metabolism

By comparing various forms of infrastructure, results from this study support a growing understanding of the biogeochemical consequences of expanded hydrologic connectivity in urban watersheds. Strong inverse relationships between DOC and NO$_3$ present across all infrastructure categories (Fig. 4c) suggest that organic carbon availability modulates inorganic nitrogen loading to streams. DOC availability has been shown to control NO$_3$ concentrations across terrestrial and aquatic ecosystems through a variety of coupled microbial processes (Hedin et al. 1998, Kaushal to streams. Additionally, the average DOC: NO$_3$ ratio, (i.e. the slope of this relationship) varied significantly across categories. Variation in this relationship is likely driven by a combination of differential N loading across categories as well as different capacities for microbial N uptake and removal.

We speculate that the location of infrastructure on the landscape may affect the relative importance of direct anthropogenic loading vs. microbial processes on DOC: NO$_3$ ratios of stream water. For instance we found high concentrations of and NO$_3$- and low DOC in streams draining septic systems. Much of this excess NO$_3$ is likely from septic plumes, but the lack of DOC may be the result of microbial C mineralization along subsurface flowpaths. On the other end of the spectrum, very low NO$_3$- and TDN in streams draining watersheds in the floodplain preservation category, which were also newly developed. In this case, the higher C:N may have been driven by lower N leakage rates as well as improved ecological function of the preserved floodplain wetlands to remove any N that does enter the groundwater from stormwater or sewage leaks.

Understanding the spatial variability in N$_2$O concentrations, as well as the processes responsible for N$_2$O production and N$_2$O removal in watersheds is useful for informing watershed management. The relationship between N$_2$O and CO$_2$ can provide insight into production mechanisms because nitrification consumes CO$_2$ while denitrification simultaneously produces N$_2$O and...
CO₂. We found a strong positive relationship between N₂O saturation and CO₂ concentrations, suggesting that denitrification was the primary source of N₂O (Figure 5c). By contrast, very low DOC: NO₃⁻ ratios (Figure 2) in stream water with highest N₂O saturation (Figure 3a) suggest that nitrification was the dominant process at these sites. Taylor and Townsend (2010) suggest that the ideal DOC: NO₃⁻ stoichiometry for denitrification is 1:1, and that persistent conditions below that are more ideal for nitrification. DOC: NO₃⁻ was consistently below 1 in streams in septic system infrastructure, suggesting that in-stream denitrification would be carbon limited. We measured DOC: NO₃⁻ consistently above 1 at sites in riparian/floodplain preservation typology, suggesting NO₃⁻ was limiting for in-stream denitrification in this infrastructure category. Conversely, the mean stoichiometric ratio was consistently near 1 in sites with inline SWM wetlands and stream burial, suggesting that denitrification may be occurring within the stream channel at these sites. While DOC: NO₃ stoichiometry in watersheds with septic systems appeared more favorable for nitrification, the positive CO₂: AOU vs. N₂O relationships in these streams suggest that these gases were produced anaerobically (by denitrification). One possible explanation for this discrepancy is that the N₂O and CO₂ observed in the stream were produced under stoichiometric conditions more favorable for denitrification along groundwater flow paths prior to emerging in the stream channel. Denitrification occurring along groundwater flowpaths may draw down the DOC concentration as it is converted to CO₂; however the initial N load in septic plumes may be too high to noticeably decline. Pabich et al. (2001) documented this phenomenon, in which DOC concentrations in a septic plume were quite high (>20 mg L⁻¹) in the upper part of the plume, and declined exponentially resulting in a very low DOC: NO₃ ratio at depth.

Overall, the relationships between CH₄ and CO₂ were much weaker and more variable than the relationships between CO₂ and N₂O (Figure 4). While CO₂ and CH₄ are sometimes correlated in wetlands and rivers with low oxygen (Richey et al. 1998), this was not the case for our study sites. Instead, CO₂ and N₂O were highly coupled, suggesting prevalence of NO₃⁻ as a terminal electron acceptor over CO₂.

### 4.3 Effects of infrastructure on N₂O Saturation and Emissions

The present study documents some of the highest N₂O concentrations currently reported in the literature for streams and rivers, ranging from 0.009 to 0.55 µM, with a median value of 0.07 µM and mean of 0.11 µM N₂O-N. This range of concentration is greater than that reported for headwater agricultural and mixed land use streams in the Midwestern United States (0.03 – 0.07 µM, Werner et al. 2012; 0.03 to 0.15 µM, Beaulieu et al. 2008). A similar range of dissolved N₂O concentrations was reported for macrophyte-rich agriculturally influenced streams in New Zealand (0.06 to 0.60µM, Wilcock and Sorrell, 2008). The only
report of higher dissolved N\textsubscript{2}O concentrations in streams is from a subtropical stream receiving irrigation runoff, livestock waste, and urban sewage (saturation ratio max of 60 compared with 47 in this study; Harrison et al. 2005).

Average daily N\textsubscript{2}O emissions were high, ranging from 5.1 to 149.6 mg N\textsubscript{2}O-N m\textsuperscript{-2} d\textsuperscript{-1}. Our values rates fall on the high end compared with numerous studies of N\textsubscript{2}O emission from urban and agriculturally influenced waterways, including agricultural drains in Japan (max= 179 mg N m\textsuperscript{-2} d\textsuperscript{-1}; Hasegawa et al. 2000) or the Humber Estuary, UK (max= 121 mg N m\textsuperscript{-2} d\textsuperscript{-1}; Barnes and Owens 1998). When the highest site (DRKV) is removed, these average daily fluxes remain high (range= 5.1 to 12.3 mg N m\textsuperscript{-2} d\textsuperscript{-1}) compared with estimates reported for nitrogen enriched agricultural and mixed land use streams in the Midwestern U.S. from Beaulieu et al. 2008 (mean = 0.84 and maximum = 6.4 mg N\textsubscript{2}O- N m\textsuperscript{-2} d\textsuperscript{-1}). Laursen and Seitzinger (2004) reported higher maximum rates (20 mg N m\textsuperscript{-2} d\textsuperscript{-1}) to our overall median N\textsubscript{2}O emission rates (13.8 mg N m\textsuperscript{-2} d\textsuperscript{-1}), and the maximum daily rates measured in tropical agricultural streams in Mexico (mean = 1.2 max= 58.8 mg N\textsubscript{2}O-N m\textsuperscript{-2} d\textsuperscript{-1}; Harrison and Matson 2003).

While our measured N\textsubscript{2}O saturation ratios were highly correlated solute concentrations and redox conditions (Table 4), emission rates sensitive to the gas transfer velocity (k_{soil}), which varied by two orders of magnitude in our study (Table 6). Correlations between TDN and N\textsubscript{2}O concentrations in WWTP influent were highly correlated to nutrient, emission rates associated with urban N loading on GHG production along urban flowpaths- which include groundwater, within pipes, and along the stream network (Tables 3 & 4). While urban streams receive a mixture of different N sources including fertilizer, wastewater, atmospheric deposition (e.g. Kaushal et al. 2011; Pennino et al. 2016), the location of aging gravity sewers adjacent to stream channels is likely to influence the relative importance of sewage on N and N\textsubscript{2}O loading to streamwater. While this source of N\textsubscript{2}O emission is likely a small portion of the global budget, gaseous losses of N can contribute significant portion of watershed-scale N budgets, which are relevant to nutrient management (Gardner et al. 2015). N\textsubscript{2}O emissions from uncollected human waste (i.e. leaky sanitary sewer lines, septic system effluent, dug pits) are largely unmeasured globally (Strokal and Kroese 2014; UNEP 2013) and warrant further study in the context of watershed management as well as local GHG accounting. Direct emissions from wastewater treatment plants (WWTPs) are well documented (Foley et al. 2010; Townsend-Small et al. 2011; Strokal and Kroese 2014; UNEP 2013), however the upstream losses of N\textsubscript{2}O from delivery pipes into streams and rivers are not (Short et al. 2014). Short et al. (2014) measured N\textsubscript{2}O concentrations in WWTP influent in Australia and determined that sanitary sewers are consistently super-saturated with N\textsubscript{2}O, with concentrations in excess of equilibrium by as much as 3.5µM. Average daily sewer pipe xN\textsubscript{2}O concentrations were 0.55 µM, which is nearly identical to the maximum xN\textsubscript{2}O measured in the present study (0.54 µM). While wastewater only contributes a portion of excess N in urban streams, further accounting for this source is necessary to improve municipal N\textsubscript{2}O budgets.
Synoptic surveys of N$_2$O saturation in Red Run and Dead Run in this study provide evidence that the entire network is a net source of N$_2$O (Fig. 5). N$_2$O saturation shows a significant decline with increasing drainage area (Table 6, Fig. 5), suggesting that emissions outpace new sources to the water column. Variability in gas concentrations headwater sites and along the 3rd order stream networks are largely explained by a combination of discharge and/or drainage area, as well as N concentrations and C:N stoichiometry in streamwater.

4.4 Effects of infrastructure on CH$_4$ Saturation and Emissions

Methane was consistently super-saturated across all streams in this study, and varied significantly across headwater infrastructure categories. The highest CH$_4$ saturation ratios were measured in sites with riparian reconnection (RRRM and RRRB) followed by streams draining in line SWM wetlands (DRKV and DRGG) (Fig. 3). As with CO$_2$, CH$_4$ saturation was negatively correlated with DO, however CH$_4$ was positively correlated with DOC: NO$_3^-$, CO$_2$, and N$_2$O, by contrast, were more strongly and positively correlated with TDN (Table 4). These patterns suggest that, along with redox conditions, carbon availability may modulate CH$_4$ production as well.

CH$_4$ concentrations in our study ranged from 0.06 to 6.08 µmol L$^{-1}$, equivalent to the mean +/- standard deviation of concentrations reported by a meta-analysis by Stanley et al. (2016). Saturation ratio (3.0 to 2157) fell within the lower range of previously measured values in agricultural streams in Canada (sat. ratio 500 to 5000, Baulch et al. 2011a). Mean daily CH$_4$ emissions estimates in this study ranged from 2.6 to 103.5 mg CH$_4$ C m$^{-2}$ d$^{-1}$ and are comparable to measurements in agricultural streams of New Zealand (Wilcock and Sorrel, 2008; 17-56 mg CH$_4$ C m$^{-2}$ d$^{-1}$) and southern Canada (20-172mg C m$^{-2}$ d$^{-1}$, Baulch et al. 2011), however these studies also measured ebullitive (i.e. bubble) fluxes, whereas the present study only examined diffusive emissions. Stanley et al. (2016) reported the average of all current CH$_4$ emission rates to be 98.7 mg C m$^{-2}$ d$^{-1}$ with a minimum of -125.3 and a maximum of 5,194 overall. While the CH$_4$ emission estimates in the present study have a large margin of uncertainty due to the nature of estimating gas flux parameters as well as the lack of ebullitive flux measurements, our sites were consistently sources to the atmosphere throughout the year at both headwater sites (Figure 3) and throughout 3rd order drainage networks (Figure 5b). Differences in CH$_4$ abundance across infrastructure categories, as well as the negative relationship between CH$_4$ saturation and TDN, suggest that CH$_4$ may increase if TDN declines with the addition of stormwater wetlands and floodplain reconnection in urban areas.
5 Conclusions

Urban watersheds are highly altered systems, with heterogeneous forms of infrastructure and water quality impairment. The present study demonstrates that N\textsubscript{2}O and CH\textsubscript{4} saturation and emissions from urbanized headwaters are on the high end of estimates currently reported in the literature. Variations in urban infrastructure (i.e. SWM wetlands, riparian connectivity, septic systems) influenced C:N stoichiometry and redox state of urban streams. These in-stream variables, along with potential direct sources from leaky sanitary sewer lines may contribute to increased GHG production and/or delivery to streams. Our results suggest that N from septic plumes and sanitary sewer lines is the principal source of N\textsubscript{2}O saturation in our study sites. Dissolved inorganic N is highly correlated with N\textsubscript{2}O in our study sites, and the highest values are only present in watersheds with aging sanitary sewer infrastructure or septic systems. Our observations of N\textsubscript{2}O saturation and emissions from urban and suburban headwater streams are comparable with streams and ditches in intensive agricultural watersheds (Harrison and Matson, 2003; Outram et al. 2012). These results suggest that streams draining medium to low-density suburban or exurban land cover are comparable to those in intensively managed agricultural areas in terms of N\textsubscript{2}O emissions.

Code availability: The authors are happy to share any and all codes used to produce this manuscript. Please contact the corresponding author with inquiries about the codes used.

Data availability: The authors have provided tables of all raw data collected for this study in the supplementary information files. These datasets will additionally be available as part of the Baltimore Ecosystem Study LTER Site archive (www.beslter.org).

Author Contributions:

R. Smith, S. Kaushal, C. Welty and M. Pennino selected sampling sites based on infrastructure typology. R. Smith, S. Kaushal and J. Beaulieu designed the gas and solute sampling design. R. Smith and J. Beaulieu analyzed samples for solute and gas concentrations respectively. C. Welty collected continuous flow data from headwater gaging stations. J. Beaulieu provided key insights into interpretation of gas concentrations and statistical analyses and gas flux estimations. M. Pennino provided data used for estimating K\textsubscript{sp}. S. Kaushal and C. Welty provided funding for the project. All coauthors provided feedback on multiple versions of the manuscript.

Competing interests: The authors declare that they have no conflict of interest.
Acknowledgements

The authors gratefully acknowledge funding from the National Science Foundation Water Sustainability and Climate program (NSF grants CBET-1058038 and CBET-1058502), as well as scientific infrastructure provided by the Baltimore Ecosystem Study LTER (www.besliter.org, NSF grant DEB-1027188). Field data collection was also partially supported by NOAA grant NA10OAR431220 to the Center for Urban Environmental Research and Education (www.cuere.umbc.edu) and the Water Resources mission area of the U.S. Geological Survey (water.usgs.gov). Daniel Jones provided advice on spatial analyses and numerous individuals including Tamara Newcomer, Tom Doody, Evan McMullen, John Urban, Shahan Haq, Julia Gorman, Julia Miller, John Kemper, Erin Stapleton, and Joshua Cole provided field assistance and/or feedback on drafts of this manuscript.

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.
References


R Core Team: R: a language and environment for statistical computing, 2014.


and subsurface flow constructed wetlands treating dairy wastewater. Journal of environmental quality 39:460–71. doi:


Werner, S.F., van Kessel, C., Browne, B.A. and Driscoll, C.T.: Three-dimensional spatial patterns of trace gas concentrations in

Wilcock, R.J. and Sorrell, B.K.: Emissions of Greenhouse Gases CH4 and N2O from Low-gradient Streams in Agriculturally


network: Implications for the effects of urban sewage and IPCC methodology. Global Change Biology 19:2999–3010. doi:

Zsolnay, A., Baigar, E., Jimenez, M., Steinweg, B., and Saccomandi, F.: Differentiating with fluorescence spectroscopy the

### Table 1. Summary of site characteristics including drainage area (km$^2$), percent impervious cover (%IC), and percent of the watershed drained by GI stormwater best management practices (%GI SWM drainage).

<table>
<thead>
<tr>
<th>Infrastructure Feature</th>
<th>Site</th>
<th>Drainage Area (km$^2$)</th>
<th>% IC Cover</th>
<th>% GI SWM Drainage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Septic Systems</td>
<td>RRSD</td>
<td>0.23</td>
<td>7.9</td>
<td>0.00</td>
<td>Low-density residential development with septic systems, minimal stormwater management with some stream burial.</td>
</tr>
<tr>
<td></td>
<td>RRSM</td>
<td>0.68</td>
<td>3.78</td>
<td>13.97</td>
<td>Suburban and commercial low-impact development converted from agriculture in early 2000s. Stormwater wetlands in upland + wide riparian buffer zones surround each stream and sanitary sewer infrastructure.</td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRM</td>
<td>0.63</td>
<td>16.4</td>
<td>100.00</td>
<td>Older suburban development (1950s) with GI located inline with stream channels, rather than dispersed across the landscape. Watershed is serviced by sanitary sewers.</td>
</tr>
<tr>
<td></td>
<td>RRRB</td>
<td>0.21</td>
<td>22.81</td>
<td>54.67</td>
<td>Older suburban and commercial development (1950s) with piped headwaters upstream of the sampling point. Watershed is serviced by sanitary sewers. No management of stormwater other than the pipe network, which also contains buried streams.</td>
</tr>
</tbody>
</table>
Table 2 Summary of results (main effects p-values) from mixed effects models examining the role of infrastructure typology and date on the following response variables: CO₂, N₂O and CH₄ saturation ratios; TDN and DOC concentrations (mg L⁻¹), BIX, and HIX (unitless).

<table>
<thead>
<tr>
<th>Main Effects</th>
<th>CO₂</th>
<th>CH₄</th>
<th>N₂O</th>
<th>TDN</th>
<th>DOC</th>
<th>BIX</th>
<th>HIX</th>
<th>DOC: NO₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure typology</td>
<td>0.496</td>
<td>0.298</td>
<td>0.488</td>
<td>0.068</td>
<td>0.200</td>
<td>0.441</td>
<td>0.020</td>
<td>&lt;0.008*</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date p-value</td>
<td>0.957</td>
<td>&lt;0.008*</td>
<td>&lt;0.008*</td>
<td>0.086</td>
<td>0.387</td>
<td>0.155</td>
<td>0.765</td>
<td>0.492</td>
</tr>
<tr>
<td>Date by Infrastructure Typology Interaction</td>
<td>&lt;0.008*</td>
<td>&lt;0.008*</td>
<td>&lt;0.008*</td>
<td>0.114</td>
<td>0.978</td>
<td>0.490</td>
<td>0.899</td>
<td>0.894</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Mean with standard error in parentheses of GHG saturation ratios, TDN and DOC concentrations (mg L\(^{-1}\)), BIX values and HIX values for each site.

<table>
<thead>
<tr>
<th>Infrastructure Typology</th>
<th>Site</th>
<th>CO(_2)</th>
<th>CH(_4)</th>
<th>N(_2)O</th>
<th>TDN</th>
<th>DOC</th>
<th>BIX</th>
<th>HIX</th>
<th>NO(_3)(^{-})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Septic Systems</td>
<td>RRSD</td>
<td>52.9 (1.1)</td>
<td>14.9 (0.5)</td>
<td>28.0 (0.7)</td>
<td>6.40 (0.20)</td>
<td>0.76 (0.12)</td>
<td>0.89 (0.02)</td>
<td>0.74 (0.01)</td>
<td>0.06 (0.01)</td>
</tr>
<tr>
<td></td>
<td>RRSM</td>
<td>13.5 (0.5)</td>
<td>25.6 (1.5)</td>
<td>5.9 (0.2)</td>
<td>3.49 (0.13)</td>
<td>1.40 (0.25)</td>
<td>0.70 (0.02)</td>
<td>0.782 (0.015)</td>
<td>0.27 (0.04)</td>
</tr>
<tr>
<td>Riparian/Floodplain Preservation</td>
<td>RRRM</td>
<td>6.6 (0.3)</td>
<td>207.3 (36.2)</td>
<td>1.7 (0.04)</td>
<td>0.59 (0.08)</td>
<td>2.89 (0.27)</td>
<td>0.67 (0.01)</td>
<td>0.85 (0.02)</td>
<td>12.16 (3.45)</td>
</tr>
<tr>
<td></td>
<td>RRRB</td>
<td>9.6 (0.4)</td>
<td>103.6 (8.6)</td>
<td>3.6 (0.1)</td>
<td>0.35 (0.02)</td>
<td>1.58 (0.18)</td>
<td>0.716 (0.01)</td>
<td>0.85 (0.01)</td>
<td>9.24 (2.43)</td>
</tr>
<tr>
<td>In Line SWM</td>
<td>DRKV</td>
<td>28.1 (1.0)</td>
<td>50.8 (8.5)</td>
<td>19.1 (0.6)</td>
<td>2.52 (0.16)</td>
<td>2.65 (0.24)</td>
<td>0.75 (0.01)</td>
<td>0.86 (0.003)</td>
<td>2.38 (0.67)</td>
</tr>
<tr>
<td></td>
<td>DRGG</td>
<td>16.3 (1.1)</td>
<td>225.8 (31.9)</td>
<td>7.9 (0.4)</td>
<td>1.16 (0.07)</td>
<td>5.32 (0.60)</td>
<td>0.73 (0.02)</td>
<td>0.83 (0.01)</td>
<td>8.72 (2.23)</td>
</tr>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td>7.9 (0.3)</td>
<td>11.3 (0.6)</td>
<td>5.1 (0.2)</td>
<td>2.68 (0.09)</td>
<td>2.64 (0.37)</td>
<td>0.81 (0.01)</td>
<td>0.83 (0.01)</td>
<td>1.42 (0.40)</td>
</tr>
<tr>
<td></td>
<td>DRIS</td>
<td>22.6 (1.0)</td>
<td>78.4 (5.8)</td>
<td>10.7 (0.5)</td>
<td>2.42 (0.09)</td>
<td>2.51 (0.27)</td>
<td>0.79 (0.01)</td>
<td>0.82 (0.01)</td>
<td>1.82 (0.44)</td>
</tr>
</tbody>
</table>
Table 4. Main effects, model coefficients, adjusted $r^2$, and overall model p-value for stepwise regression models examining the relationship between continuous variables and GHG saturation ratios. The model coefficient is the main effect of each parameter, and the absolute value of this coefficient signifies the relative contribution of each predictor. A * indicates the predictor with the greatest influence for each response variable (CO$_2$, CH$_4$ and N$_2$O). Rows with n.a. indicate that the predictor variable was not retained in the final model.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>CO$_2$</th>
<th>CH$_4$</th>
<th>N$_2$O</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDN</td>
<td>1.08*</td>
<td>n.a.</td>
<td>1.10*</td>
</tr>
<tr>
<td>Temperature</td>
<td>-0.22</td>
<td>0.25</td>
<td>-0.26</td>
</tr>
<tr>
<td>DO</td>
<td>-0.46</td>
<td>-0.27</td>
<td>-0.37</td>
</tr>
<tr>
<td>HIX</td>
<td>0.09</td>
<td>-0.15</td>
<td>0.13</td>
</tr>
<tr>
<td>BIX</td>
<td>0.11</td>
<td>n.a.</td>
<td>0.15</td>
</tr>
<tr>
<td>%IC</td>
<td>n.a.</td>
<td>-0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>%SWM</td>
<td>0.18</td>
<td>0.16</td>
<td>0.31</td>
</tr>
<tr>
<td>log(DOC:NO$_3$)</td>
<td>0.32</td>
<td>0.55*</td>
<td>0.19</td>
</tr>
</tbody>
</table>

**Overall Model Fit**

<table>
<thead>
<tr>
<th>Adjusted $r^2$</th>
<th>CO$_2$</th>
<th>CH$_4$</th>
<th>N$_2$O</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.78</td>
<td>0.5</td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P-value</th>
<th>CO$_2$</th>
<th>CH$_4$</th>
<th>N$_2$O</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt;0.0008^*$</td>
<td>$&lt;0.0008^*$</td>
<td>$&lt;0.0008^*$</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Summary of gas flux estimations for the four sites with continuous flow data. Average, standard error (s.e.), and number of measurements (n) are listed for CO\textsubscript{2} (g C m\textsuperscript{-2} day\textsuperscript{-1}), CH\textsubscript{4} (mg C m\textsuperscript{-2} day\textsuperscript{-1}), N\textsubscript{2}O (mg N m\textsuperscript{-2} day\textsuperscript{-1}), and predicted \(k_{600} (m\ day^{-1})\).

<table>
<thead>
<tr>
<th>Infrastructure typology</th>
<th>Site</th>
<th>Parameter</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>s.e.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td>CO\textsubscript{2} g C m\textsuperscript{-2} d\textsuperscript{-1}</td>
<td>2.37</td>
<td>23.12</td>
<td>11.51</td>
<td>6.12</td>
<td>5</td>
</tr>
<tr>
<td>Inline SWM</td>
<td>DRGG</td>
<td>CH\textsubscript{4} mg C m\textsuperscript{-2} d\textsuperscript{-1}</td>
<td>53.28</td>
<td>548.01</td>
<td>134.55</td>
<td>30.18</td>
<td>2</td>
</tr>
<tr>
<td>Inline SWM</td>
<td>DRKV</td>
<td></td>
<td>3.39</td>
<td>23.81</td>
<td>10.30</td>
<td>1.74</td>
<td>18</td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td></td>
<td>0.61</td>
<td>5.51</td>
<td>2.55</td>
<td>1.00</td>
<td>18</td>
</tr>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td>N\textsubscript{2}O mg N m\textsuperscript{-2} d\textsuperscript{-1}</td>
<td>0.61</td>
<td>5.51</td>
<td>2.55</td>
<td>1.00</td>
<td>18</td>
</tr>
<tr>
<td>Inline SWM</td>
<td>DRGG</td>
<td></td>
<td>2.27</td>
<td>1339.62</td>
<td>102.51</td>
<td>75.57</td>
<td>3</td>
</tr>
<tr>
<td>Inline SWM</td>
<td>DRKV</td>
<td></td>
<td>3.26</td>
<td>62.98</td>
<td>16.80</td>
<td>5.29</td>
<td>18</td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td></td>
<td>2.19</td>
<td>12.11</td>
<td>6.69</td>
<td>2.19</td>
<td>11</td>
</tr>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td></td>
<td>2.13</td>
<td>24.21</td>
<td>12.33</td>
<td>6.43</td>
<td>5</td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRGG</td>
<td>N\textsubscript{2}O mg N m\textsuperscript{-2} d\textsuperscript{-1}</td>
<td>60.45</td>
<td>565.17</td>
<td>149.63</td>
<td>33.91</td>
<td>3</td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRKV</td>
<td></td>
<td>1.90</td>
<td>8.61</td>
<td>5.14</td>
<td>0.79</td>
<td>18</td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td></td>
<td>2.57</td>
<td>16.98</td>
<td>7.03</td>
<td>2.63</td>
<td>11</td>
</tr>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td>(k_{600} (m\ day^{-1}))</td>
<td>3.84</td>
<td>19.20</td>
<td>10.97</td>
<td>4.47</td>
<td>5</td>
</tr>
<tr>
<td>Inline SWM</td>
<td>DRGG</td>
<td></td>
<td>12.82</td>
<td>122.59</td>
<td>28.02</td>
<td>7.06</td>
<td>3</td>
</tr>
<tr>
<td>Inline SWM</td>
<td>DRKV</td>
<td></td>
<td>2.40</td>
<td>8.89</td>
<td>5.39</td>
<td>0.73</td>
<td>18</td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td></td>
<td>2.57</td>
<td>13.91</td>
<td>6.45</td>
<td>2.33</td>
<td>11</td>
</tr>
</tbody>
</table>
Table 6. Covariates and model fit parameters for linear models describing drivers of gas saturation ratios (CO₂, CH₄ and N₂O) from longitudinal surveys of Dead Run and Red Run. ‘X’ s denote that a given parameter was used in the final model while dashes (-) denote parameters not used.

<table>
<thead>
<tr>
<th>Covariates Tested</th>
<th>CO₂ Sat. Ratio</th>
<th>CH₄ Sat. Ratio</th>
<th>N₂O Sat. Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>log of drainage area (km²)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Watershed (Dead Run vs. Red Run)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Season</td>
<td>-</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DOC (mg L⁻¹)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DIC (mg L⁻¹)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TN (mg L⁻¹)</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Log of Q (m³ s⁻¹)</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Location (tributary vs. main stem)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DOC:TN molar ratio</td>
<td>X</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>TN x log of drainage area interaction</td>
<td>X</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>DOC x log of drainage area interaction</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Model AIC**

<table>
<thead>
<tr>
<th>CO₂ Sat. Ratio</th>
<th>CH₄ Sat. Ratio</th>
<th>N₂O Sat. Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>336.85</td>
<td>542.14</td>
<td>263.59</td>
</tr>
</tbody>
</table>

**Overall model r²**

<table>
<thead>
<tr>
<th>CO₂ Sat. Ratio</th>
<th>CH₄ Sat. Ratio</th>
<th>N₂O Sat. Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.789</td>
<td>0.153</td>
<td>0.795</td>
</tr>
</tbody>
</table>

**Overall Model p-value**

<table>
<thead>
<tr>
<th>CO₂ Sat. Ratio</th>
<th>CH₄ Sat. Ratio</th>
<th>N₂O Sat. Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.008</td>
<td>0.0082</td>
<td>&lt;0.008</td>
</tr>
</tbody>
</table>
Figure 1: Site map of headwater stream sites within Red Run and Dead Run watersheds. Green stars signify bi-weekly sampling sites, and black dots signify longitudinal sampling points sampled seasonally. Land cover categories are colored based on the National Land Cover Database, with dark red areas signifying dense urban land cover, light red signifying medium urban land cover, and green colors signifying forested or undeveloped areas. Close-up views of Dead Run and Red Run on the right represent the study watersheds, with areas that are captured by stormwater management structures (detention basins, wetlands, sand filters, etc.) shaded in gray.
Figure 2: Boxplot of molar DOC: NO$_3^-$ ratio across sites in watersheds with differing infrastructure typologies. The median of each dataset is signified by the middle horizontal line for each category. Boxes signify the range between first and third quartiles (25$^{th}$ and 75$^{th}$ percentiles). Vertical lines extend to the minimum and maximum points in the dataset that are within 1.5 times the inter-quartile range. Points signify data that fall above or below this range. Letters represent significant (p <0.01) differences between infrastructure typologies for DOC: NO$_3^-$ across all sampling dates, determined using a linear mixed effects model.
Figure 3: Boxplot of CO$_2$, CH$_4$, and N$_2$O saturation ratios across stream sites in varying infrastructure categories. Letters denote significant pairwise differences across streams for a given gas from linear mixed effects models with ‘watershed’ as a main effect. Boxes signify the range between first and third quartiles (25th and 75th percentiles). Vertical lines extend to the minimum and maximum points in the dataset that are within 1.5 times the interquartile range. Points signify outliers outside of 1.5 times the interquartile range.
Figure 4: Scatterplots of a) N$_2$O saturation vs. xsCO$_2$-AOU (µM), CH$_4$ saturation vs. anaerobic CO$_2$, and c) relationships between NO$_3^-$ and DOC. Lines denote significant (p<0.01) correlations among gas or solute concentrations, which vary by infrastructure category.
Figure 5: Longitudinal variability in CO$_2$ (a-b), CH$_4$ (c-d), and N$_2$O (e-f) saturation ratios from spring and fall synoptic surveys of Dead Run and Red Run. Dotted lines denote tributaries to each watershed, while straight lines denote the main stem sites.
GHG emissions from streams in agricultural watersheds have been investigated in numerous studies, but less is known about streams draining urban watersheds.
GHG emissions from streams in agricultural watersheds have been investigated in numerous studies, but less is known about streams draining urban watersheds.
GHG emissions from streams in agricultural watersheds have been investigated in numerous studies, but less is known about streams draining urban watersheds.

Streams and rivers are globally significant sources of nitrous oxide (N2O), carbon dioxide (CO2), and methane (CH4) (e.g., Seitzinger et al. 2000; Beaulieu et al. 2011; Bastviken et al. 2011; Raymond et al. 2013). The interactive effects of climate and land cover change have increased greenhouse gas emissions (GHG) from streams and rivers by altering the biogeochemical controls of ecosystem metabolism (i.e., nutrient stoichiometry, organic matter quality, redox state, and temperature), (e.g. Kaushal et al. 2014a; Beaulieu et al. 2009; Dinsmore et al. 2009; Baulch et al. 2011; Harrison and Matson 2003). Urban stormwater and sanitary sewer infrastructure – including stormwater wetlands, stream burial in pipes, gravity sanitary sewer lines, and septic systems – influences nutrient loading (Shields et al. 2008; Kaushal and Belt 2012; Newcomer et al. 2012; Pennino et al. 2014; Beaulieu et al. 2015) and may have implications for GHG production as well. Numerous studies have examined the role of point sources of nutrients such as wastewater treatment plant (WWTP) effluent on urban N2O emissions (Foley et al. 2010; Townsend-Small et al. 2011; Strokal and Kroeze 2014; Beaulieu et al. 2010), but few have examined the role of nonpoint source nutrient loading on N2O emissions from urban streams. The nonpoint source N loads from gravity sewers and septic systems, however, may contribute substantially to urban N2O emissions (Beaulieu et al. 2010; Short et al. 2014). Aquatic N2O production and emissions have been linked to microbial transformations of excess N loading, as well as reduced oxygen availability (Beaulieu et al. 2011; Rosamond et al. 2012). While stormwater-control wetlands and other forms of green infrastructure (GI) may reduce N2O production in streams by reducing excess N inputs, GI may increase both N2O and CH4 inputs to streams and groundwater due to CH4 and N2O production that occurs within the GI unit (Søvik et al. 2006; VanderZaag et al. 2010). Despite considerable funds spent on restoring aging infrastructure and improving water quality in cities globally (Doyle et al. 2008), the role of urban water infrastructure on biogeochemical cycles and GHG production is a major source of uncertainty.

The International Panel on Climate Change (IPCC) includes N2O emissions from agricultural, but not urban streams, in the global N2O inventory based on nitrogen inputs from fertilizer and manure (Nevison...
N loading to streams can be as high in urban as in agricultural watersheds, but the relationship between N and N2O emissions may differ substantially in urban and agricultural watersheds. Some key differences include: 1) the source and quantity of anthropogenic N loading to streams, 2) the C:N ratio of stream water and groundwater, and 3) the degree to which surface and groundwater flowpaths are altered by infrastructure. These factors are likely to be influenced by stormwater and sanitary sewer infrastructure designs (Søvik et al. 2006; Collins et al. 2010; Kaushal et al. 2011). Stormwater management may promote anoxic conditions and increase C:N ratio of stream water if wetlands are created along the urban watershed continuum (e.g. Søvik et al. 2006; Newcomer et al. 2012). Stormwater management can reduce C:N ratios, if streams are buried in storm drains (Elmore and Kaushal 2008; Pennino et al. 2016; Beaulieu et al. 2014). Sanitary sewer infrastructure may additionally contribute to GHG emissions from urban streams by direct leakage of gases or excess nitrogen from sewer lines (Yu et al. 2013; Short et al. 2014).

Inverse relationships between dissolved organic carbon (DOC) and nitrate (NO3-) concentrations persist across a wide variety of ecosystems ranging from soils to streams to oceans (e.g., Aitkenhead-Peterson and McDowell 2000; Dodds et al. 2004; Kaushal and Lewis 2005; Taylor and Townsend 2010). Recently, inverse relationships between DOC and NO3- have also been reported for urban environments from groundwater to streams to river networks (Mayer et al. 2010; Kaushal and Belt 2012; Kaushal et al. 2014a).

A suite of competing biotic process may control this relationship, by either: 1) assimilating or reducing NO3- in the presence of bioavailable DOC, or 2) producing NO3- regardless of DOC status (Hedin et al. 1998; Dodds et al. 2004; Kaushal and Lewis 2005; Taylor and Townsend 2010). The former category includes heterotrophic denitrification, which oxidizes organic carbon to CO2 and reduces NO3- to N2O + N2 (Knowles, 1982), and assimilation of inorganic N (Wymore et al. 2015; Caraco et al. 1998; Kaushal and Lewis 2005). In the second category, nitrification is a chemoautotrophic process that produces NO3- by oxidizing NH4+, and consumes CO2. Nitrification also yields N2O as an intermediate product, and has been shown to dominate N cycling processes in low-DOC environments (Taylor and Townsend, 2010). In urban watersheds, denitrification is often limited by DOC due to increased N loading and/or decreased connectivity with carbon-rich soils in the riparian zone (Mayer et al. 2010; Newcomer et al. 2012). The interactive effects of increased anthropogenic C and N loading and biogeochemical transformations have the potential to alter GHG production and emissions from streams (Kaushal et al. 2014b).

The goal of the present study was to identify patterns and potential drivers related to GHG dynamics in urban headwater streams draining different forms of infrastructure (stream burial, septic systems, in-line SWM wetlands and riparian/floodplain preservation). Although less considered, GHG emissions may be an unintended consequence of urban water quality impairments and biogeochemical processes occurring within and downstream of urban infrastructure. An improved understanding of the relationship between infrastructure type and biogeochemical functions along the urban watershed continuum is critical for minimizing unintended consequences of water quality management (Kaushal and Belt 2012). Additionally,
a better understanding of the contribution of urban watersheds to global GHG emissions will be critical, given that urbanization is the fastest form of land-use change and urban areas contain greater than 60% of Earth’s population (Foley et al. 2005; Bellucci et al. 2012; Ciais et al. 2013).

that are part of the Baltimore Long-Term Ecological Research (LTER) project (www.besler.org) were sampled every two weeks for water chemistry and dissolved gases. Sampling sites were located in the Red Run and Dead Run subwatersheds of the Gwynns Falls that were developed at different times (Fig. 1). Previous work in the Baltimore LTER project has extensively characterized the hydrology, biogeochemistry, and geomorphology of the Gwynns Falls stream network (e.g., Doheny 1999; Groffman et al. 2004, Nelson et al. 2006; Kaushal et al. 2008, Shields et al. 2008, Meierdiercks et al. 2010; Ryan et al. 2010; Sivirichi et al. 2011, Newcomer et al. 2012; Newcomer Johnson et al. 2014; Pennino et al. 2014; Pennino et al. 2016; Bhaskar et al. 2012, 2015).

Study sites were selected based on differences in stormwater and sanitary sewer infrastructure within each of eight headwater watersheds. Dead Run (15 km²) and Red Run (17 km²) are both dominated by medium to high-density residential and commercial land. Dead Run was developed between the 1950s and 1970s. We selected eight headwater stream watersheds, each of which drained one of four distinct infrastructure typologies. These typologies were based on having similar land cover, development age, stormwater infrastructure design, and sanitary infrastructure. A comprehensive description of attributes in each typology can be found in Table 1, however for simplicity we have abbreviated the typologies based on the dominant infrastructure feature as follows: 1) stream burial, 2) in-line stormwater management (SWM) wetlands, 3) riparian/floodplain preservation, and 4) septic systems.

with channelized or buried streams as part of the stormwater infrastructure and aging sanitary sewer lines that are often cracked and leaking to the subsurface.

Stormwater wetlands and ponds drain a portion of the Dead Run watershed and are located in-line with stream channels. In contrast, Red Run was intensively developed in the 2000s and stormwater infrastructure reflects more infiltration-based designs such as stream buffer zones, A few small areas with low-density development built in the 1960s that are served by septic systems are located in the northern part of Red Run
A few small areas with low-density development built in the 1960s that are served by septic systems are located in the northern part of Red Run.

and three blank samples were taken at each field site.

and three blank samples were taken at each field site.

and three blank samples were taken at each field site.

along the two main paired watersheds of
along the two main paired watersheds of

We used data from these surveys to determine whether or not headwater streams have more variable GHG saturation than the higher order parts of the stream network. Reach-scale hydrologic mass balances were calculated along the main stem of Red Run and Dead Run from these synoptic surveys following methods detailed previously (Kaushal et al. 2014a, Newcomer Johnson et al. 2014). Along each reach of the main stem, relative contributions of inflow were calculated following Eq. (1):

\[ Q_{GW} = Q_{DS} - Q_{US} - Q_{TRIB}, \]  

(1)

where \( Q_{GW} \) is the net groundwater input, estimated by difference using field measurements of \( Q_{DS}, Q_{US}, \) and \( Q_{TRIB}. \) \( Q_{DS} \) is discharge measured in the main stem (m³ s⁻¹) at the bottom of a reach, \( Q_{US} \) is discharge in the main stem at the top of a reach, and \( Q_{TRIB} \) is inflow from major tributaries.
For example, t
production or consumption of biogenic gases. However, the biological influence on dissolved gas concentration can be confounded with the physical effects of gas exchange. To correct for the effect of different gas exchange rates across streams, we calculated the ratios between apparent oxygen utilization (AOU) and \( x_s \)CO\(_2\). AOU is calculated as the difference between \( O_2 \) concentration at equilibrium with the atmosphere and measured dissolved oxygen in the stream. Positive values of AOU therefore signify net consumption of \( O_2 \) along watershed flowpaths, and negative AOU values signify net production \( O_2 \). Under aerobic conditions, respiration of organic matter consumes \( O_2 \) and produces \( CO_2 \) in approximately a 1:1 molar ratio (Schlessinger 1997). Therefore, 1 mole of AOU should result in 1 mol of \( x_s \)CO\(_2\). This ratio was then used, with an offset to 1.2:1 to account for differences in diffusion constants for the two gases (Stumm and Morgan 1981; Richey et al. 1988), to determine the proportion of \( CO_2 \) produced from aerobic respiration. For instance, 1 mol of AOU would result in 1 mol of \( CO_2 \) excess if aerobic respiration where the only \( CO_2 \) source. A \( CO_2 \) excess value greater than 1 mol would be indicative of other \( CO_2 \) sources, namely anaerobic respiration, which produces \( CO_2 \) without consuming \( O_2 \). This framework was used to calculate the percentage of \( CO_2 \) produced from anaerobic vs. abiotic processes. Anaerobic \( CO_2 \) concentrations were calculated as the difference between aerobically produced \( CO_2 \) (assumed equivalent to AOU) and measured \( CO_2 \) concentration.

Gas emissions were calculated using Eq. (5), in which, and FGHG is the flux (g m\(^{-2}\) d\(^{-1}\)) of a given gas at ambient temperature, \( d \) is water depth (m), and KGT (day\(^{-1}\)), is the air-water gas exchange rate for a given gas at ambient temperature as in Eq. (4)

\[
F = d \times (\text{[}] - \text{[} e \text{]}),
\]

The air-water gas exchange rate was estimated for each site and sampling date using an energy dissipation model (Tsivoglou and Neal 1976). This model describes \( K20 \) as a function of water velocity (\( V \), m day\(^{-1}\)), water surface gradient (\( S \)), and a site-specific constant called the escape coefficient (\( C_{esc} \), m\(^{-1}\)) (Eq. 6).

\[
K20 = S \times V,
\]

(6)

We estimated \( S \) at each GHG sampling site by measuring the change in elevation over a reach
Cesc is a parameter related to additional factors other than streambed slope and velocity that affect gas-exchange including streambed roughness and the relative abundance of pools and riffles. We estimated Cesc for our sampling sites using measurements of sulfur hexafluoride (SF\(_6\)) gas exchange rate (K_{SF6}) from 15 tracer injection experiments carried out across a range of flow conditions in four streams within 5 km\(^2\) of our study sites (Pennino et al. 2014).

Cesc was calculated to be 0.198 m\(^{-1}\) (n=15, r\(^2\)=0.81, P= 5.48 x10\(^{-6}\)). The 95% confidence interval of this Cesc based on measured K\(_{20}\) values was ±0.058 which corresponds to ±29% of a given gas flux estimate. This estimate of Cesc from these nearby sites was assumed to be representative of the 8 stream reaches investigated in this study. The uncertainty associated with Cesc was small compared to the difference in estimated flux across sites. Areal flux data was thus interpreted with caution, and only examined in terms of the magnitude across all sites and in comparisons with literature values.

We converted K_{SF6} to K for CO2, CH4, and N2O by multiplying K_{SF6} by the ratio of Schmidt numbers for SF6 and each measured gas (Stumm and Morgan 1981). K was also adjusted to 20°C (K\(_{20}\)) following Eq. (7), where KT is K for a given gas at ambient temperature

\( KT = K_{20} \cdot \frac{T}{298} \)

\( T \) is the ambient temperature in Kelvin and \( K_{20} \) is the gas exchange coefficient at 20°C. The Schmidt number is given by the ratio of the kinematic viscosity \( \nu \) to the diffusivity \( D \) of the gas in water.

\( S = \frac{\nu}{D} \)

The kinematic viscosity \( \nu \) and the diffusivity \( D \) of gases in water are temperature dependent and can be expressed as functions of temperature in Kelvin, \( T \), using the following equations:

\( \nu = \nu_0 \cdot e^{\frac{-E_v}{RT}} \)

\( D = D_0 \cdot e^{\frac{-E_d}{RT}} \)

Where \( \nu_0 \) and \( D_0 \) are the kinematic viscosity and diffusivity at a reference temperature, \( R \) is the gas constant, and \( E_v \) and \( E_d \) are the activation energies for viscosity and diffusivity, respectively.

The Schmidt number for CO2 is 1.3, for CH4 is 0.7, and for N2O is 0.5. These values are used to calculate the Schmidt number for each gas by dividing the gas's Schmidt number by the Schmidt number for SF6.

\( S_{gas} = \frac{S_{SF6}}{S_{gas}} \)

The gas exchange coefficient for each gas at ambient temperature is then calculated using the Schmidt number and the gas exchange coefficient for SF6 at 20°C.

\( K_{gas} = K_{SF6} \cdot \frac{S_{SF6}}{S_{gas}} \)

This approach provides a consistent method for converting gas exchange coefficients across different gases and temperatures.
3.2 Effects of urban infrastructure on dissolved organic matter quality
3.2 Effects of urban infrastructure on dissolved organic matter quality

Very high N\textsubscript{2}O saturation ratios were measured in headwaters of both Red Run and Dead Run, which were not representative of the remainder of the drainage network (Fig. 5).

Instead, a logarithmic decline was observed between the sites with highest N\textsubscript{2}O saturation and the main stem along hydrologic flowpaths from engineered headwaters to larger order streams. Headwater CH\textsubscript{4} saturation ratios were not markedly different from that in the main stem.

Greenhouse gas emissions varied substantially across sites and dates. The magnitude of CO\textsubscript{2}, CH\textsubscript{4}, and N\textsubscript{2}O emissions increased with discharge due to the dependence of K\textsubscript{20} on slope and velocity.

Emissions during three high-flow sampling dates (over 0.015 m\textsuperscript{3} s\textsuperscript{-1} for all sites) increased the variance of overall mean gas emission rates estimates.

When these high emission rates were removed, average daily CO\textsubscript{2} emissions (± standard error) was twenty to 100-fold higher at DRKV 39.5 (±15.5) g C m\textsuperscript{-2} day\textsuperscript{-1} than the other sites, due in part to the tenfold high stream surface slope at DRKV.
runoff,

rates

rates

rates

rates

rates

rates

rates

rates

rates

rates

rates

rates

rates
N₂O emissions from agricultural runoff are currently included in IPCC estimates, but emissions associated with urban ecosystems are not currently accounted for (Ciais et al. 2013). Urban and agricultural streams are similar in that they receive excess nitrogen inputs from the watershed, including N inputs from contaminated groundwater. Key differences arise when considering N₂O budgets, however. Whereas agricultural stream emissions are estimated based on annual fertilizer inputs, N in urban streams is derived from diffuse, spatially heterogeneous nonpoint sources. For instance, studies in Baltimore have found that atmospheric deposition and human waste contribute approximately 25% and 50% of nitrate inputs, while the remainder is derived from soils and plant materials (Kaushal et al. 2011; Pennino et al. 2016). The proportion of these sources and others is likely to vary widely across and within watersheds.

Recent reviews have suggested that
are similar in that they receive excess nitrogen inputs from the watershed, including N inputs from contaminated groundwater. Key differences arise when considering N\textsubscript{2}O budgets, however. Whereas agricultural stream emissions are estimated based on annual fertilizer inputs, N in urban streams is derived from diffuse, spatially heterogeneous nonpoint sources. For instance, studies in Baltimore have found that atmospheric deposition and human waste contribute approximately 25 % and 50 % of nitrate inputs, while the remainder is derived from soils and plant materials (Kaushal et al. 2011; Pennino et al. 2016). The proportion of these sources and others is likely to vary widely across and within watersheds.

Recent reviews have suggested that

N\textsubscript{2}O emissions from agricultural runoff are currently included in IPCC estimates, but emissions associated with urban ecosystems are not currently accounted for (Ciais et al. 2013). Urban and agricultural streams are similar in that they receive excess nitrogen inputs from the watershed, including N inputs from contaminated groundwater. Key differences arise when considering N\textsubscript{2}O budgets, however. Whereas agricultural stream emissions are estimated based on annual fertilizer inputs, N in urban streams is derived from diffuse, spatially heterogeneous nonpoint sources. For instance, studies in Baltimore have found that atmospheric deposition and human waste contribute approximately 25 % and 50 % of nitrate inputs, while the remainder is derived from soils and plant materials (Kaushal et al. 2011; Pennino et al. 2016). The proportion of these sources and others is likely to vary widely across and within watersheds.

Recent reviews have suggested that
atmospheric deposition and human waste contribute approximately 25% and 50% of nitrate inputs, while the remainder is derived from soils and plant materials (Kaushal et al. 2011; Pennino et al. 2016). The proportion of these sources and others is likely to vary widely across and within watersheds. Recent reviews have suggested that
Measurements of CH$_4$

Measurements of CH$_4$

Measurements of CH$_4$

Measurements of CH$_4$

Measurements of CH$_4$
Methane concentrations were consistent with prior studies, showing that streams are commonly supersaturated with CH4 (e.g. Jones and Mulholland 1998; Wilcock and Sorrel 2008; Baulch et al. 2011; Werner et al. 2012). In contrast with IPCC methodology (Ciais et al. 2013), there is growing evidence that human impacts on watersheds influence CH4 emissions from streams (Kaushal et al. 2014b, Crawford and Stanley 2015; Stanley et al. 2015). Prior studies have found that CH4 production tends to be elevated in streams with fine benthic sediments, an influx of organic matter, or significant wetland drainage (Dinsmore et al. 2009; Dawson et al. 2002; Baulch et al. 2011). Significant negative relationships between TDN and CH4
were detected in this study, and elevated CH4 concentrations in streams draining intact floodplains and/or stormwater management wetlands.

An increasing number of scientific studies have compiled GHG budgets of anthropogenic and ecological emissions across cities (e.g., Brady and Fath, 2008; Hoornweg et al. 2011; Weisert et al. 2014). Understanding both the anthropogenic and ecological components of a regional GHG budget is crucial for setting GHG targets and managing ecosystem services (Bellucci et al. 2012). The role of human activities on GHG emissions from agriculturally impacted waterways is well recognized (Ciais et al. 2013; Nevison 2000). However, further studies examining the magnitude and variations in GHG emissions along the urban watershed continuum, which explicitly includes flowpaths from engineered infrastructure to streams and rivers (e.g. Kaushal and Belt 2012), are necessary. As cities and populations continue to expand globally, GHG emissions from wastewater are likely to rise. A greater understanding of the interplay between urban water infrastructure and biogeochemical processes is necessary to mitigate negative consequences of N2O, CH4, and CO2.

M. Pennino provided feedback on multiple versions of the manuscript

<table>
<thead>
<tr>
<th>Infrastructure typology</th>
<th>Site</th>
<th>Gas</th>
<th>Average</th>
<th>s.e.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td></td>
<td>0.50</td>
<td>0.18</td>
<td>17</td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRGG</td>
<td>CO2</td>
<td>0.64</td>
<td>0.10</td>
<td>15</td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRKV</td>
<td></td>
<td>39.55</td>
<td>8.72</td>
<td>19</td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td></td>
<td>2.04</td>
<td>0.48</td>
<td>17</td>
</tr>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td>N2O</td>
<td>0.57</td>
<td>0.22</td>
<td>17</td>
</tr>
<tr>
<td>Category</td>
<td>Code</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRGG</td>
<td>0.53</td>
<td>0.09</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRKV</td>
<td>44.07</td>
<td>9.79</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td>1.01</td>
<td>0.23</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Stream Burial</td>
<td>DRAL</td>
<td>2.29</td>
<td>0.75</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRGG</td>
<td>1.67</td>
<td>0.33</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>In-line SWM</td>
<td>DRKV</td>
<td>57.56</td>
<td>9.63</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Floodplain Preservation</td>
<td>RRRB</td>
<td>9.36</td>
<td>1.67</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>