What role does stream restoration play in nutrient management?

Roderick W. Lammers & Brian P. Bledsoe

To cite this article: Roderick W. Lammers & Brian P. Bledsoe (2017): What role does stream restoration play in nutrient management?, Critical Reviews in Environmental Science and Technology, DOI: 10.1080/10643389.2017.1318618

To link to this article: http://dx.doi.org/10.1080/10643389.2017.1318618
What role does stream restoration play in nutrient management?

Roderick W. Lammers\textsuperscript{a} and Brian P. Bledsoe\textsuperscript{b}

\textsuperscript{a}Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado, USA; \textsuperscript{b}College of Engineering, University of Georgia, Athens, Georgia, USA

\section*{ABSTRACT}
Nutrient pollution is a pervasive water quality problem. Stream restoration has been proposed as a novel approach to reduce loading and increase nutrient processing within streams. We summarize evidence from the literature on the efficacy of stream restoration for reducing nutrient loading and increasing nutrient removal in stream ecosystems. We also analyze published data on streambank phosphorus concentrations and riparian and stream denitrification rates to improve understanding of the potential benefits of stream restoration for phosphorus retention and nitrogen removal. Finally, we discuss the role of stream restoration in nutrient management and provide recommendations for practice and future research.

\section*{1. Introduction}
Eutrophication of aquatic systems—excessive nutrient concentrations and subsequent accelerated primary production—is a pressing water quality problem in the United States and around the world (Smith et al., 1999; Smith, 2003). Nitrogen and phosphorus, nutrients required for plant and animal growth, are often limiting in these ecosystems (Elser et al., 2007), meaning that small increases in the availability of one or both of these compounds can lead to large increases in biomass. Although necessary for proper ecological function, excessive nitrogen and/or phosphorus can lead to algal blooms, depleting dissolved oxygen (potentially lethal to aquatic life), producing harmful toxins, impairing aesthetics and recreation, and fouling water infrastructure. This eutrophication significantly degrades water quality, harming aquatic organisms and beneficial human use of water resources. According to the US Environmental Protection Agency (EPA), nutrient pollution is the third largest source of water quality impairment in rivers and streams and the second largest for lakes, ponds, and reservoirs (EPA, 2015).
Significant effort has been made to improve water quality by reducing nutrient loading to streams from upland sources; however, in-stream nutrient removal and retention may be significant (Peterson et al., 2001; Reisinger et al., 2015), leading to the suggestion that enhancing natural removal and sequestration mechanisms may be a viable strategy to improve water quality. Stream restoration is a booming enterprise in the United States (>$1 billion per year industry), and the majority of these projects list water quality improvement as one of their objectives (Bernhardt et al., 2005). Despite the rapid growth of stream restoration as a management tool over the last two decades, there is a notably limited set of studies that examine the efficacy of stream restoration for nutrient removal, processing, and retention.

The purpose of this review is to synthesize existing evidence for the ability of stream restoration to reduce nutrient loads and concentrations in aquatic systems. Furthermore, we provide a quantitative analysis of streambank phosphorus concentrations and stream and riparian denitrification rates—issues of concern for both nutrient loading and removal. We focus on denitrification as a permanent removal pathway for nitrogen. Phosphorus has no analogous removal processes, but stream restoration may decrease phosphorus inputs by reducing streambank erosion. This analysis will aid in future quantification of nutrient-related benefits of restoration, and may support development of crediting programs, which can spur additional investment in restoration as a nutrient management strategy.

2. Background

2.1. Phosphorus

The chemical partitioning of phosphorus is important for understanding its transport. Phosphorus species are relatively insoluble and are typically adsorbed to soil particles. They have a high affinity for the large specific surface area of clay and silt particles and are also found bound in various metal oxyhydroxides including Fe-OH, Al-OH, and Ca-OH (Brady and Weil, 2002). The partitioning of phosphorus among its various states determines its bioavailability for uptake by organisms, which is directly tied to its importance as a limiting nutrient. The relative abundance of bioavailable phosphorus in stream sediment has been shown to vary markedly within single study sites (1–55%; Veihe et al., 2011) and between study areas (averaging 0.5–22% of total phosphorus; Nellesen et al., 2011; Howe et al., 2011; Hubbard et al., 2003; McDowell and Sharpley, 2001; McDowell and Wilcock, 2007; Thompson and McFarland, 2007).

Phosphorus speciation and subsequent bioavailability is a dynamic process; therefore, there may be a lag from when phosphorus is introduced to a stream and when the effects of this loading are manifested (Meals et al., 2010). Additionally, sediment may serve as either a sink or a source of phosphorus depending on whether the sediment sorptive capacity is higher or lower than the in-stream dissolved phosphorus concentrations (e.g. Hoffman et al., 2009; McDaniel et al.,
The bioavailability of phosphorus has important implications for its effects on water quality. However, most water quality monitoring programs focus only on total phosphorus because of the difficulty in both quantifying bioavailability and predicting how the forms of phosphorus will change over time. Unlike denitrification of nitrate, there is no biotic or abiotic process that effectively removes phosphorus from an ecosystem. Therefore, phosphorus “removal” is likely only temporary biotic uptake, although burial and storage in floodplain or lacustrine sediment may be a longer-term removal mechanism (Records et al., 2016).

Recent evidence has made it increasingly clear that bank and bed erosion may be a significant source of particulate phosphorus loading to streams (Fox et al., 2016), typically accounting for between 10% (Sekely et al., 2002) and 40% (Howe et al., 2011) of the total phosphorus load in an individual watershed. However, phosphorus and sediment loading rates from bank erosion are highly variable both within and among watersheds, ranging from <1 to 1,000s of kg/km-year (Table 1). Nitrogen loading from bank erosion is commonly neglected (but see Walter et al., 2007), likely because phosphorus is more commonly found adsorbed to soil particles than nitrogen, and most nitrogen studies focus on dissolved, inorganic forms. Despite increasing recognition of the role of bank erosion as a phosphorus source, quantifying this contribution remains difficult, in part because of uncertainty in bank phosphorus concentrations. We addressed this uncertainty by analyzing available data on bank phosphorus content to help practitioners quantify the magnitude of this potentially potent nutrient source.

### Table 1. Summary of bank sediment and phosphorus loading rates from select studies.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Location</th>
<th>Bank P concentrations (mg/kg)</th>
<th>Sediment loading rate (tonne/km-year)</th>
<th>Phosphorus loading rate (kg/km-year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeWolfe et al. (2004)</td>
<td>Vermont, USA</td>
<td>369–740</td>
<td>10–1,333</td>
<td>10–840</td>
</tr>
<tr>
<td>Hubbard et al. (2003)</td>
<td>Mississippi, USA</td>
<td>168–237</td>
<td>7,574–17,604</td>
<td>1,170–2,719</td>
</tr>
<tr>
<td>Kronvang et al. (2012b)</td>
<td>Funen, Denmark</td>
<td>400–1,400</td>
<td>NA</td>
<td>4.2–167</td>
</tr>
<tr>
<td>Langendoen et al. (2012)</td>
<td>Vermont, USA</td>
<td>422–973</td>
<td>0.1–3,266</td>
<td>0.1–2,970</td>
</tr>
<tr>
<td>Miller et al. (2014)</td>
<td>Oklahoma, USA</td>
<td>239–427</td>
<td>300–12,857</td>
<td>320–4,104</td>
</tr>
<tr>
<td>Nellesen et al. (2011)</td>
<td>Iowa, USA</td>
<td>184–347</td>
<td>NA</td>
<td>3.4–20.7</td>
</tr>
<tr>
<td>Peacher (2011)</td>
<td>Missouri, USA</td>
<td>197–637</td>
<td>9.2–447.5</td>
<td>2.9–114</td>
</tr>
<tr>
<td>Thoma et al. (2005)</td>
<td>Minnesota, USA</td>
<td>249–452</td>
<td>9,147</td>
<td>3,589</td>
</tr>
<tr>
<td>Zaimes et al. (2008b)</td>
<td>Iowa, USA</td>
<td>303–555</td>
<td>5–304</td>
<td>2–123</td>
</tr>
</tbody>
</table>

* Tonne is a metric ton (1,000 kg).

* Reported nitrogen concentrations of 400–2,100 mg/kg and nitrogen loading rates of 447–6,113 kg/km-year.

In aquatic systems, nitrogen cycles between organic and inorganic forms through plant uptake, organic material decomposition, mineralization, and microbially mediated nitrification (formation of nitrate from ammonia) and denitrification. Denitrification is the anaerobic reduction of nitrate by heterotrophic bacteria...
under anoxic conditions, leading to the production of N₂O or N₂ gas (Hill, 1996). Denitrification completely removes nitrate from a system, making it a desirable target for restoration as opposed to biotic assimilation, which only temporarily alters nitrogen availability (e.g. plants utilize nitrogen but release it back to the stream when they die and decompose). Denitrification rates are highest in saturated soils and streambeds (water providing a source of dissolved nitrate and anoxic conditions) and when organic carbon is available (to serve as an energy source). However, denitrification is rarely a constant process. High rates of denitrification may occur in discrete locations (“hot spots”) and at discrete points in time (“hot moments”) when conditions are right (McClain et al., 2003). However, it remains unclear just how significant these hot spots and hot moments are for controlling ecosystem function (Bernhardt et al., 2017).

Encouraging denitrification through stream restoration is desirable; however, denitrification rates may remain high even in impacted streams. Forested and urban riparian soils have similar denitrification potential (Groffman and Crawford, 2003), suggesting that under the right conditions, namely high water tables, riparian denitrification will occur even in developed areas. On the other hand, suburban and agricultural riparian areas may have lower organic carbon availability compared with forested sites, a potential constraint on denitrification (Watson et al., 2010). Stream restoration may also fail to encourage denitrification, and rates may not be significantly different between restored and unrestored reaches (Groffman et al., 2005; Harrison et al., 2012; Tuttle et al., 2014). In fact, restored reaches may have lower denitrification rates due to coarser streambed sediment (Weigelhofer et al., 2013) and lower organic carbon availability (Gift et al., 2010), side effects of construction activities. Focusing on denitrification for nitrogen removal is commendable, but it is a difficult goal to achieve in practice. We conducted an analysis of in-stream and riparian denitrification rates to help determine what controls this important nitrogen removal process and to provide guidance for restoration.

### 2.3. Stream restoration

Stream or river restoration is a broad topic with varying definitions (National Research Council, 1992; Society for Ecological Restoration, 2004; Wohl et al., 2005). Studies included in this review fit the definition of stream restoration broadly by incorporating modifications to streams, riparian areas, and/or the watershed with the intent of providing some ecological and/or human benefit. The studies incorporated within this review use a variety of stream restoration strategies, including Natural Channel Design (NCD) (Rosgen, 1996, 2006). NCD attempts to restore a degraded reach to match the geomorphic form of some nearby reference reach that is considered representative of natural and stable conditions. This approach has been criticized for failing to account for differences in watershed condition and processes (e.g. hydrology and sediment flux) between degraded and reference sites and the focus on stability, rather than ecological
improvement, as a restoration goal (Simon et al., 2007; Lave, 2009). Nutrient removal and retention in aquatic systems is a dynamic process and therefore restoration that focuses specifically on relevant biotic and abiotic processes is likely to be the most successful (Wohl et al., 2005, 2015).

While the most commonly stated objective of stream restoration projects is water quality management, fewer than 10% of projects have any kind of post-restoration monitoring (Bernhardt et al., 2005). Even fewer projects have published data on changes in water quality, making it difficult to assess whether these projects achieved their stated goals. Recent reviews have assessed stream restoration success broadly (Palmer et al., 2014; Newcomer Johnson et al., 2016) and with a focus on specific processes (e.g. denitrification in restored hyporheic zones; Merill and Tonjes, 2014). However, to date, there has been no comprehensive assessment of the ability of specific restoration strategies to improve nutrient retention and processing. The purpose of this review is to synthesize the current state of the science on stream restoration as a strategy for nutrient removal and retention.

3. Methods

3.1. Streambank phosphorus concentrations

A number of studies have quantified streambank phosphorus concentrations in diverse locations. These studies focused on a variety of aspects of bank erosion and phosphorus loading including the influence of adjacent land use; variability between stream reaches and between bed, bank, and upland areas; relative abundance of various forms of phosphorus; and quantification of total watershed phosphorus loading from bank erosion. However, there has been no large-scale analysis of the available bank phosphorus data to determine whether observed trends are consistent between studied geographic areas. We collected bank phosphorus data from the literature and supplemented this data set with our own field data from Big Dry Creek, Westminster, CO and Lick Creek, Durham, NC. We analyzed our bank samples for both total phosphorus (EPA Method 3051a) and bioavailable phosphorus (Mehlich-3).

Bank sampling and laboratory analysis of phosphorus content is expensive. We assessed the feasibility of using regional soil phosphorus data as a more cost-effective approach to estimate bank phosphorus concentrations. We obtained coarse-scale soil phosphorus data for the conterminous United States from a recent US Geological Survey publication (Smith et al., 2013). We compared bank total phosphorus concentrations from various studies to upland soil (surface, A, and C horizons) phosphorus concentrations within a 50-mile radius (chosen to yield ~10 upland data points). For some studies, bank phosphorus data were separated by study stream if they were geographically distant. Weighted least squares regression was used to relate mean bank phosphorus concentration to upland phosphorus concentrations for each location, incorporating a pseudo $R^2$ value to account for the applied weights (Willett and Singer, 1988). Each point was weighted by the inverse of its squared
residual from ordinary least squares regression (giving points closer the line higher weight). We quantified uncertainty using 1,000 bootstrap replicates with data resampling (with replacement) for each site, yielding new mean concentrations with which regression was performed. This allowed for the quantification of a nonparametric 95% confidence interval (CI) for the pseudo $R^2$ values.

Phosphorus bioavailability is important for understanding its effect on water quality, but it remains an elusive parameter to quantify. We assessed variability in streambank phosphorus bioavailability using data from eight sources. In each case, Mehlich-3 phosphorus was assumed to represent bioavailable phosphorus. Although various extraction methods may be used to estimate bioavailable phosphorus (e.g. water extraction), Mehlich-3 is the most widely used and therefore provided the largest data set.

Individual laboratory methods differ in their ability to extract adsorbed phosphorus from soils, leading to potentially different results and complicating direct comparison of data from various sites (Kleinman et al., 2001). Analytical methods for determining total phosphorus concentrations of bank soils were divided into four major categories to test their relative influence on observed data. The *aqua regia* approach utilizes a combination of nitric and hydrochloric acid in conjunction with microwave heating (Crosland et al., 1995; U.S. EPA, 2007). Others use only hydrochloric acid, also with heat (Pardo et al., 2004). Alternative methods include sulfuric acid digestion (U.S. EPA, 1983; method 365.4 colorimetric, automated) and alkaline oxidation (Dick and Tabatabai, 1977).

### 3.2. Denitrification rates

Reviewers have examined denitrification in freshwater and marine ecosystems (Seitzinger, 1988) and riparian zones (Martin et al., 1999), and have even attempted to quantify effects of various controlling variables on denitrification rates (Saunders and Kalf, 2001; Piña-Ochoa and Álvarez-Cobelas, 2006). However, numerous studies on stream and riparian denitrification have been published since these reviews were conducted, and no review has collected stream and riparian denitrification data for concurrent analysis.

Data on streambed and riparian denitrification rates were obtained from 98 peer-reviewed studies (69 stream and 38 riparian with some overlap, Table S1). Actual denitrification rates were collected from text, tables, and figures. If potential denitrification rates (i.e. measured rates after adding abundant nitrate and carbon) were available, these data were obtained as well (Table S2). Information on quantification approach, including specific method, spatial scale (i.e. core or reach), temporal scale (i.e. days, months, or years), and whether the analysis was performed in the laboratory or field was collected. Other variables of interest include dominant watershed land use, riparian vegetation type, Köppen-Geiger climate classification (Kottek et al., 2006), and average streamwater nitrate concentration. Since riparian
nitrate levels were not consistently reported (e.g. soil concentration versus groundwater concentration), these data were excluded from the analysis.

The collected data are meant to be representative rates from streams and riparian zones. We either used mean rates reported by the authors or we averaged point measurements across space and time for each site. Different methods measure rates at different scales. In streams, reach-scale rates can be measured directly using tracers; however, many stream data (and all riparian data) are obtained as point measurements and averaged across the sample area to obtain denitrification rates on a per-area basis. This analysis accounts for differences in method and scale, and direct comparison of riparian and stream denitrification rates should be appropriate.

Denitrification in natural systems is notoriously difficult to measure. A variety of methods have been developed; however, they all have their drawbacks (Groffman et al., 2006). Details on the benefits and shortcomings of various methods are provided by Groffman et al. (2006), and therefore will not be repeated here. However, we do include a brief description of the major methods used by studies included in this analysis (Table 2).

We used several approaches to quantify the influence of selected variables on denitrification rates: linear regression, regression tree, and random forest. Regression trees are similar to linear regression techniques in that the response variable can be predicted based on a set of predictor variables. However, while linear regression treats the entire predictor space the same, tree-based approaches stratify the predictor space into regions. Instead of a linear regression model, the output is a “tree” with binary splits based on predictor variable values or categories. A random forest approach extends this idea by creating a large number of regression trees and averaging the result. We assessed variable importance by these three methods using R (version 3.2.2; R Core Team, 2015). We used the “rpart” package (Therneau et al., 2015) for regression tree analysis and the “party” package

---

**Table 2. Summary of common denitrification methods.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylene inhibition</td>
<td>Acetylene (C₂H₂) blocks the conversion of nitrous oxide (N₂O) to dinitrogen gas (N₂), leaving the easier-to-measure N₂O as the terminal product of denitrification. This method is inexpensive and lends itself to large numbers of samples. However, it may underestimate actual denitrification rates.</td>
</tr>
<tr>
<td>¹⁵N tracer</td>
<td>Use of a nitrogen isotope tracer allows for quantification of denitrification by tracking isotopically labeled nitrogen gas production. Complications with this method may arise if adequate mixing of the added tracer is not achieved.</td>
</tr>
<tr>
<td>Direct N₂ or N₂:Ar measurement</td>
<td>Direct measurement of N₂ produced from denitrification is difficult because of the high background concentrations in the atmosphere. However, it is more suitable in aquatic systems where gas exchange is limited. Use of mass spectrometry to quantify denitrification via the ratio of N₂:Ar in produced gas is also increasing in use.</td>
</tr>
<tr>
<td>Mass balance</td>
<td>Quantitative nitrogen mass balance can allow for estimation of denitrification in aquatic systems. To be effective, all other nitrogen fluxes must be quantified, with the discrepancy being attributed to denitrification. This is obviously prone to significant error.</td>
</tr>
</tbody>
</table>

See Groffman et al. (2006) for a more detailed description.
(Hothorn et al., 2006; Strobl et al., 2007, 2008) for random forest. Each of these approaches yielded variable importance metrics, which were normalized to their sum for comparison between methods. The average of these normalized metrics was used as an overall indicator of variable importance. Climate types were consolidated by main climate group (e.g. tropical, temperate, arid) if there were fewer than five observations. The exception was the sole study in a polar climate (Gooseff et al., 2004), which was left independent. We retained this data point because removing it did not significantly influence results.

3.3. Restoration assessment

Different restoration practices target different nutrient removal pathways and are therefore likely to have different levels of impact on nutrient loading in stream and riparian ecosystems. We performed simple Monte Carlo simulations to compute the mean nutrient removal potential, along with associated uncertainty, of four restoration strategies: bank stabilization, floodplain reconnection, riparian buffer restoration, and hyporheic zone restoration. Simulations assumed a conceptual 1-km long, 10-m wide restored reach with a 20-m buffer on one bank and a 20-m reconnected floodplain on the other. Bank and bed stabilization and floodplain reconnection were analyzed for phosphorus removal (avoided loading from erosion and increased deposition on floodplains). We used representative bank erosion rates (Table 1), bank phosphorus concentrations (Fig. 1), and floodplain phosphorus retention rates (Kronvang et al., 2007). Hyporheic zone restoration, riparian buffer restoration, and floodplain reconnection were assessed for their ability to remove nitrate via denitrification. We used stream and riparian denitrification rates collected from the literature (Section 3.2).

4. Results

4.1. Streambank phosphorus concentrations

Figure 1 shows the variability in bank total phosphorus concentrations reported in 17 studies from around the United States and the world (overall mean 337 mg/kg, 5th percentile 72.5 mg/kg, 95th percentile 707 mg/kg). This variability can lead to significant uncertainty in estimated phosphorus loading rates from bank erosion. Streambank phosphorus concentrations are well correlated with upland surface soil data (Fig. 2; slope = 0.505; $R^2 = 0.45$, 95% CI: 0.14–0.58). Regression results using data from the A and C horizons have higher slopes but weaker relationships (data not shown; A: slope = 0.674; $R^2 = 0.34$, 95% CI: 0.06–0.54; C: slope = 0.878; $R^2 = 0.31$, 95% CI: 0.07–0.49). Dominant land use of the upland data does not appear to have a significant influence on the regression results.

Bioavailable phosphorus accounts for, on average, 11.7% of total streambank phosphorus in the available studies (Fig. 3). These data are variable but suggest bioavailable phosphorus is usually less than one quarter of total phosphorus (90%
Percent bioavailability is negatively correlated with total phosphorus (TP; Spearman $\rho = -0.20$, $p < 1 \times 10^{-6}$) suggesting that the bioavailable phosphorus fraction decreases as TP increases, and it may be inappropriate to assume a constant percent bioavailability for any value of TP. However, this negative correlation is relatively weak, and more data are needed at TP values greater than $\sim 600$ mg/kg to yield more definitive results.

Statistically significant differences were observed between different phosphorus extraction methods based on a nonparametric, rank-based pairwise comparison (Siegel and Castellan, 1988) (Fig. 4). However, these differences pale in comparison to interstudy variability. Multiple linear regression results suggest that differences between studies explain an order of magnitude greater variability in TP concentrations than the analytical method.

### 4.2. Streambed and riparian denitrification rates

Areal denitrification rates in stream and riparian systems are shown in Fig. 5. A two-sided Wilcoxon rank-sum test suggests that stream denitrification rates are significantly different than riparian rates at the 95% confidence level ($p = 0.016$). However, the difference in the medians of these two data sets is only 0.84 mg N m$^{-2}$ h$^{-1}$ (1.85 and 1.01 mg N m$^{-2}$ h$^{-1}$ for stream and riparian, respectively). Given the several orders of magnitude range in both data sets, this difference may not be practically meaningful.
Streamwater nitrate concentration is generally considered one of the most significant controls on stream denitrification rates. Our compiled data set shows a similarly strong relationship between these two variables (Fig. 6). A fitted power regression has a larger exponent than previous studies (Table 3; 0.72 compared to 0.51, Mulholland et al., 2009; 0.44, Bohlke et al., 2009; and 0.62, Roley et al., 2012). However, using just data collected using a $^{15}$N tracer method, our calculated exponent (0.56) is more similar to those found by others using $^{15}$N tracer data (Mulholland et al., 2009). Fitted exponents for data collected using the other methods are all substantially higher (0.71–0.75), suggesting a greater increase in denitrification rate per unit increase in nitrate concentration.

**Figure 2.** (A) Locations of studies used in this analysis (Bledsoe et al., 2000; Hubbard et al., 2003; McDowell et al., 2003; Walter et al., 2007; Thompson and McFarland, 2007; Schilling et al., 2009; Hongthanat, 2010; Tufekcioglu, 2010; Howe et al., 2011; Nellesen et al., 2011; Peacher, 2011; Miller et al., 2014) and unpublished data from Big Dry Creek, Westminster, CO and Lick Creek, Durham, NC. (B) Mean bank phosphorus concentrations versus mean surface soil (top 5 cm) phosphorus concentrations. Solid line is a weighted least squares linear regression [Bank P [mg/kg] = 15.6 + 0.505×Upland P [mg/kg]]. Error bars are the interquartile range of individual points. Gray lines are bootstrap-estimated regression lines, indicating the range of uncertainty. Points are differentiated based on major land use type from the upland data set.
The Michaelis-Menten equation is commonly used to model the saturation effects of enzyme-substrate reactions:

\[
\text{Den} = \frac{(\text{NO}_3^- \cdot V_{\text{max}})}{(\text{NO}_3^- + K_s)}
\]

Where \(\text{Den}\) is the denitrification rate [mg N m\(^{-2}\) h\(^{-1}\)], \(\text{NO}_3^-\) is the stream nitrate concentration [\(\mu\)g L\(^{-1}\)], \(V_{\text{max}}\) is the maximum obtainable value of \(\text{Den}\) [mg N m\(^{-2}\) h\(^{-1}\)], and \(K_s\) is the nitrate concentration at which \(\text{Den}\) is half of \(V_{\text{max}}\) [\(\mu\)g L\(^{-1}\)]. Fitted values of \(V_{\text{max}}\) and \(K_s\) were 31.8 mg N m\(^{-2}\) h\(^{-1}\) and 3,466 \(\mu\)g L\(^{-1}\), an order of magnitude higher than those reported by Mulholland et al. (2009) (3.9 and 422, respectively) but only 2–3 times higher than those of Böhlke et al. (2009). Our fitted values were more similar to these previous studies when we used only \(^{15}\)N tracer data from our compiled data set (Table 3). The power regression model appears to fit the data better (Akaike information criterion, AIC = 405) than the Michaelis-Menten model (AIC = 1,708). This corresponds with previous conclusions that a saturation-law model may not be appropriate when comparing denitrification data across streams (Böhlke et al., 2009; Mulholland et al., 2009).

Actual denitrification as a percent of denitrification potential was significantly different for riparian versus stream data (medians = 15.6% and 3.5%, respectively; two-sided Wilcoxon signed-rank test \(p = 0.014\)). Given that streams overall tend to have slightly higher denitrification rates, this suggests that streams have greater

---

**Figure 3.** Percent of total phosphorus that is bioavailable (Mehlich-3 extraction) from a variety of sources. Box plots indicate the median, quartiles, and nonoutlier maximum and minimums (1.5 times the interquartile range). Open circles are outliers. The dashed horizontal line is the mean of the composite data set, and the dotted horizontal lines are the 95th and 5th percentiles. [A] Hon-gthanat (2010); [B] Howe et al. (2011); [C] Hubbard et al. (2003); [D] McDowell et al. (2003); [E] Miller et al. (2014); [F] Nellesen et al. (2011); [G] unpublished data from Big Dry Creek, Westminster, CO and [H] Lick Creek, Durham, NC.
Denitrification potential than riparian areas, although a smaller proportion of this potential is being utilized. Denitrification potential and rates were reported on either a per-mass or per-area basis. The percentages reported above incorporate both data sets, but additional analysis required separating these two data sets. There was a positive correlation between percent of potential and per-mass actual rates (Spearman $r = 0.64$, $p = 0.012$, $n = 15$) but no correlation with per-area actual rates (Spearman $r = -0.09$, $p = 0.783$, $n = 12$). The reason for this discrepancy is unclear. However, there is a greater frequency of low potential percentages for the per-mass data set, which could affect results. The first analysis suggests that as actual denitrification rates increase, they are approaching the potential of the system. However, the second analysis suggests that percent of potential is relatively consistent across systems regardless of the \textit{in situ} denitrification rate. More data are needed to analyze these relationships in more detail.

The variable importance analysis (Fig. 7) suggests that stream nitrate concentration is the dominant predictor of stream denitrification rates. However, land use, method, and climate are also important. Agricultural and urban streams had higher rates than reference systems. Temperate and continental climates without noticeable dry seasons tended to have higher denitrification than other climate types. For the riparian data set, climate was by far the most important predictor, followed by method, temporal scale, and riparian vegetation type. Climate types Cfa (temperate, no dry season, hot summer) and Dfb (continental, no dry season, warm summer) had the highest denitrification rates. However, these results may

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Comparison of bank total phosphorus concentration by extraction method. Box plots indicate the median, quartiles, and nonoutlier maximum and minimums (1.5 times the interquartile range). Open circles are outliers. Brackets show differences in medians and $p$-values from a rank-based pairwise comparison (Siegel and Castellan, 1988). See Section 3.1 for details on laboratory methods.}
\end{figure}
be somewhat skewed because only agricultural and urban sites (no reference sites) were represented in these climate types. Bare riparian buffers tended to have slightly higher denitrification than vegetated types (grass and forested buffers did not have significantly different rates). The results from the three methods (linear regression, regression tree, and random forest) agreed well, giving us confidence in the results.

Figure 5. Summary of published denitrification rates in riparian and stream ecosystems. Box plots indicate the median, quartiles, and non-outlier maximum and minimums (1.5 times the interquartile range). Open circles are outliers.

Figure 6. Stream denitrification rate versus in-stream nitrate concentration. Solid line is a power regression equation fit to the data \( \log_{10}(\text{Den}) = -1.91 + 0.72 \log_{10}(\text{Nitrate}) \) or, alternatively, \( \text{Den} = 0.012 \times \text{Nitrate}^{0.72}, \ R^2 = 0.45 \). Dash-dot line is fitted Michaelis-Menten model \( (V_{\text{max}} = 31.8 \text{ mg N m}^{-2} \text{ h}^{-1}, K_s = 3,466 \text{ mg L}^{-1}) \). Points are differentiated to indicate differences in denitrification quantification methods.
The influence of method on denitrification rate was examined in more detail using a Tukey-adjusted pairwise comparison of the means for the log-transformed denitrification data. For the stream data, the two most common methods, $^{15}$N tracers and acetylene block, were not significantly different ($p = 0.89$), neither were the mass balance and N$_2$ flux methods ($p = 0.99$). However, all remaining differences were significant, with mass balance and N$_2$ flux yielding higher denitrification rates than $^{15}$N tracer and acetylene block results (all $p$-values $< 0.05$). Higher rates from mass balance results are expected because this method accounts for all forms of nitrate uptake, not just denitrification. This could also explain why our Michaelis-Menten parameters were higher than others reported in the literature; the mass balance data may skew the results. N$_2$ flux rates are not expected to be higher since this method only accounts for net denitrification (gross denitrification minus nitrogen fixation, Groffman et al., 2006). Differences among methods were not statistically significant for the riparian data set, despite method appearing as the second most important explanatory variable in the variable importance analysis (Fig. 7).

### 4.3. Restoration assessment

Our simple simulations quantified the relative efficacy of various stream restoration practices for nutrient reduction and removal (Fig. 8). Bank stabilization had much

<table>
<thead>
<tr>
<th>Table 3. Summary of fitted models for denitrification rate v. nitrate for this and previous studies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model parameters</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Den[mg N m$^{-2}$ h$^{-1}$] = a$^{\text{NO}_3^-}$ [µg L$^{-1}$]$^b$</td>
</tr>
<tr>
<td>a = 0.012; b = 0.62</td>
</tr>
<tr>
<td>a = 0.01; b = 0.51</td>
</tr>
<tr>
<td>NA</td>
</tr>
<tr>
<td>a = 0.036; b = 0.71</td>
</tr>
</tbody>
</table>

Michaelis-Menten relationship $\text{Den} = \frac{(\text{NO}_3^- + V_{\text{max}})/(\text{NO}_3^- + K_s)}{V_{\text{max}}}.$

<table>
<thead>
<tr>
<th>$V_{\text{max}}$ [mg N m$^{-2}$ h$^{-1}$]; $K_s$ [µg L$^{-1}$]</th>
<th>$V_{\text{max}}$ [µmol N m$^{-2}$ h$^{-1}$]; $K_s$ [µmol L$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{\text{max}} = 31.8; K_s = 3,466$</td>
<td>$V_{\text{max}} = 640; K_s = 180$</td>
</tr>
<tr>
<td>$V_{\text{max}} = 3.9; K_s = 422$</td>
<td>$V_{\text{max}} = 2,274; K_s = 248$</td>
</tr>
<tr>
<td>NA</td>
<td>$V_{\text{max}} = 483; K_s = 69^*$</td>
</tr>
</tbody>
</table>

Various units were used in development of power relationships. While coefficients (a) are not directly comparable between units, exponents (b) are. Michaelis-Menten models for the acetylene block, N$_2$ flux, and mass balance methods were not statistically significant. $^*$Coefficient not statistically significant at $p < 0.05$. The influence of method on denitrification rate was examined in more detail using a Tukey-adjusted pairwise comparison of the means for the log-transformed denitrification data. For the stream data, the two most common methods, $^{15}$N tracers and acetylene block, were not significantly different ($p = 0.89$), neither were the mass balance and N$_2$ flux methods ($p = 0.99$). However, all remaining differences were significant, with mass balance and N$_2$ flux yielding higher denitrification rates than $^{15}$N tracer and acetylene block results (all $p$-values $< 0.05$). Higher rates from mass balance results are expected because this method accounts for all forms of nitrate uptake, not just denitrification. This could also explain why our Michaelis-Menten parameters were higher than others reported in the literature; the mass balance data may skew the results. N$_2$ flux rates are not expected to be higher since this method only accounts for net denitrification (gross denitrification minus nitrogen fixation, Groffman et al., 2006). Differences among methods were not statistically significant for the riparian data set, despite method appearing as the second most important explanatory variable in the variable importance analysis (Fig. 7).

### 4.3. Restoration assessment

Our simple simulations quantified the relative efficacy of various stream restoration practices for nutrient reduction and removal (Fig. 8). Bank stabilization had much
Figure 7. Average normalized importance metric from the regression tree, random forest, and multiple linear regression results for the stream and riparian denitrification rate data sets.

Figure 8. Results of Monte Carlo simulations of nutrient load reduction potential for (A) phosphorus and (B) nitrogen. Box plots indicate the median, quartiles, and nonoutlier maximum and minimums (1.5 times the interquartile range). Circles are outliers (offset to more clearly show the data).
greater potential to remove phosphorus from the system (609 ± 1,175 kg P/km-year; mean ± sd) compared to floodplain reconnection (25 ± 15 kg P/km-year). Riparian restoration had the greatest nitrogen removal potential (1,086 ± 4,973 kg N/km-year), followed by hyporheic restoration (226 ± 692 kg N/km-year) and floodplain reconnection (86 ± 402 kg N/km-year). High removal potential for riparian restoration is largely a result of larger surface area compared to the streambed. These simple calculations assume the maximum possible benefits of restoration (e.g. complete bank stabilization, no denitrification prior to restoration). The goal is to provide estimates of relative removal potential of various strategies, and it is inappropriate to interpret these results more broadly.

5. Discussion

5.1. Stream restoration for phosphorus reduction

5.1.1. Phosphorus and bank erosion

As discussed previously, bank erosion in unstable streams can be a significant source of particulate phosphorus. Empirical evidence suggests that bank stabilization (e.g. bioengineering, bendway weirs, rip-rap), often in conjunction with other restoration techniques, can significantly reduce pollutant loading (Meals, 2001; Carline and Walsh, 2007). However, post-restoration monitoring of a stream where bank stabilization and grade control structures were installed showed no change in nutrient concentrations (Selvakumar et al., 2010). Bank erosion is clearly an important phosphorus source in some watersheds (Table 1), but quantifying potential benefits of bank stabilization for phosphorous retention is difficult.

Estimating historic bank erosion rates is only a first step to assessing the potential for bank stabilization to reduce phosphorus loading since these rates may not be representative of future channel evolution. Conceptual channel evolution models predict that eroding channels will at some point adjust to a new, stable state (Schumm et al., 1984; Simon, 1989; Hawley et al., 2012; Cluer and Thorne, 2014; Booth and Fischenich, 2015), assuming no further disturbances to the sediment and hydrologic regimes. Therefore, current or historic erosion rates cannot simply be extrapolated for actively evolving channels as these systems should eventually stabilize on their own; although, the time scale may range from years to decades or more. In addition, deposition of eroded sediment on banks and floodplains may be a significant sink in some watersheds (Hupp et al., 2013; Kronvang et al., 2013), meaning loading estimates from bank erosion alone without accounting for sediment storage processes may overestimate total export.

An additional source of uncertainty in estimating loading rates is variability in bank phosphorus content. Bank phosphorus concentrations are associated with a number of factors (soil texture, pH, metals concentrations, and the geology of the soil parent material; Brady and Weil, 2002) and are often highly variable, even in a single reach (e.g. Bledsoe et al., 2000; Nellesen et al., 2011; Schilling et al., 2009). Streambank phosphorus concentrations
may also be higher in intensively farmed catchments (Palmer-Felgate et al., 2009) or in deforested areas (Haggard et al., 2007), although others have shown little correlation to land use (Zaimes et al., 2008b; Tufekcioglu, 2010; Nellesen et al., 2011).

In some cases, streambank phosphorus concentrations have been found to be positively correlated with silt-clay content (Cooper and Gilliam, 1987; Bledsoe et al., 2000; Agudelo et al., 2011; Kerr et al., 2011; Young et al., 2012, 2013), while others have found no correlation (Schilling et al., 2009; Hongthanat, 2010; Howe et al., 2011; Veihe et al., 2011). The reason for this discrepancy is likely rooted in differences in the types of clays in each of these regions. Phosphorus binds strongly to iron and aluminum oxides in soil. Therefore, soils with high concentrations of these compounds (especially in amorphous form) will likely have higher sorbed TP concentrations. These oxides are common in Ultisols and Oxisols, while Mollicsols and other soil types dominated by silicate clays have low phosphorus binding capabilities (Brady and Weil, 2002, p. 617). Studies in areas where silicate clays are dominant do not show a correlation between silt-clay content and TP (Schilling et al., 2009; Hongthanat, 2010), while those with high oxide clays do demonstrate this positive correlation (Bledsoe et al., 2000; Young et al., 2012). In these cases, silt-clay content may just be a proxy for metal oxide abundance. Several studies found positive correlations between metal oxide concentrations and TP, independent of the silt-clay:TP relationship (Hongthanat, 2010; Kerr et al., 2011; Young et al., 2012).

While direct quantification of bank phosphorus content through field sampling and analysis is the best approach, coarse estimates may be obtained based on regional soil phosphorus data (Fig. 2). There appears to be a strong correlation between these two data sets, although uncertainty analysis indicates high variability. Points on Fig. 2 fall below the 1:1 line, indicating that while the two data sets are correlated, upland phosphorus concentrations tend to be higher than those in streambanks. While estimating actual streambank phosphorus concentrations based on this relationship is not advisable, this correlation can at least be used to assess relative bank phosphorus levels in a preliminary assessment of phosphorus loading potential.

Differences in laboratory analysis are generally thought to have a significant influence on soil phosphorus data; however, our results suggest that differences between common extraction methods for quantifying TP of soils have a smaller effect than may have been previously thought. Nevertheless, it is still advisable to consider differences in laboratory techniques when comparing soil phosphorus data between studies. Additionally, using alternative extraction methods to estimate bioavailable phosphorus (e.g. Mehlich-3 or water extraction) is recommended to better understand potential impacts of phosphorus loading on water quality (Fox et al., 2016). Bioavailable phosphorus may be only ~10% of the total (Fig. 3), meaning that phosphorus loading estimates from bank erosion could differ by an order of magnitude, depending on the type of phosphorus considered.
5.1.2. Phosphorus and other restoration techniques

Other restoration techniques, including riparian and floodplain restoration, also influence phosphorus dynamics. Phosphorus retention in riparian buffers is complex. Although buffers may trap particulate and dissolved phosphate in surface runoff (Lee et al., 2003; Newbold et al., 2010), groundwater phosphate concentrations often increase through buffers (Peterjohn and Correll, 1984; Spruill, 2000), potentially due to phosphate release from the reduction of iron oxyhydroxides in soils (Jordan et al., 1993). These offsetting processes can result in little or no net retention of phosphorus in buffers (Newbold et al., 2010). Similarly, restored floodplains can be net phosphorus sources, even if they are effective nitrogen sinks (Jones et al., 2015). However, the type of restored floodplain may have an impact. A restored wetland-channel complex and two-stage ditch, a popular form of floodplain reconnection in agricultural areas, both reduced phosphorus export (Richardson et al., 2011; Davis et al., 2015; Mahl et al., 2015).

Even though riparian buffers may be net phosphorus sources, they can reduce phosphorus loading indirectly by altering stream dynamics. Vegetated stream-banks tend to have lower erosion rates than nonvegetated banks (Smith, 1976; Simon and Collison, 2002; Miller et al., 2014). Additionally, riparian buffers may be significant sources of in-stream wood and organic carbon to streams (Stanley et al., 2012), both of which can increase in-stream nutrient uptake (Bernhardt and Likens, 2002; Roberts et al., 2007).

5.2. Stream restoration and nitrogen

Stream restoration typically targets nitrogen removal by increasing lateral and vertical hydrologic connectivity to encourage biochemical nitrogen removal, typically via denitrification. Hydrologic connectivity can be lost in degraded streams for a number of reasons. For example, channel incision lowers riparian groundwater tables, reducing denitrification potential by disconnecting the relatively anoxic, nitrate-laden groundwater with available organic carbon in the upper soil horizons. This incision results in higher nitrate loading to streams (Schilling et al., 2006) and higher in-stream nitrate concentrations (Groffman et al., 2002; Böhlke et al., 2007), but not in all cases (Schilling and Jacobson, 2014). Channel incision can also reduce overbank flooding, disconnecting the stream from its floodplain, an important nutrient sink. Degraded streams often become geomorphically simple, losing the bedform complexity and bed sediment heterogeneity that drive vertical hydrologic connectivity with the hyporheic zone. The hyporheic zone may be small, but it is a vitally important area for biochemical processing and nutrient removal (Hester and Gooseff, 2010). Strategies to restore hydrologic connectivity and provide conditions suitable to denitrification have the greatest potential for addressing nitrogen loading.
5.2.1. Nitrogen and lateral connectivity

Riparian buffers have long been recognized for their potential to remove pollutants in surface and subsurface flow before they reach the stream. Under the right conditions, properly functioning riparian buffers (typically consisting of intact forest) can remove >90% of inflowing nitrate (Peterjohn and Correll, 1984; Jordan et al., 1993; Pinay et al., 1993; Spruill, 2000; Liu et al., 2014). A comprehensive review of published nitrate removal rates found an average effectiveness of 74.2 ± 4.0% (mean ± SE, Mayer et al., 2005). The majority of this removal takes place in the subsurface (Peterjohn and Correll, 1984; Hill, 1996) where conditions amenable for denitrification are found (low dissolved oxygen and available organic carbon).

Generally, nutrient removal effectiveness increases with buffer width and age (Mayer et al., 2007; Orzetti et al., 2010; King et al., 2016). The consensus seems to be that wider is better, although just how wide is up for debate. Effective buffer width is dependent upon a number of site-specific factors including slope, soil type, incoming nutrient load, and buffer age. As buffers mature, root depth, root density, and organic carbon availability are expected to increase, leading to more conducive conditions for denitrification.

Floodplain reconnection promotes similar removal mechanisms as riparian buffers; encouraging subsurface denitrification in the riparian zone while retaining particulate and dissolved nutrients in surface flow. Floodplain reconnection aims to restore optimal conditions for denitrification and nutrient retention by raising the channel bed, removing or breaching levees, and/or re-grading and lowering banks. Natural floodplains have been demonstrated to be important areas for nitrate removal (e.g. Forshay and Stanley, 2005), and reconnected floodplains have the potential to be effective nutrient sinks (Valett et al., 2005; Fink and Mitsch, 2007), removing small but potentially significant (0.6–23.7%) portions of a river’s nutrient load (Sheibley et al., 2006; Kronvang et al., 2007; Roley et al., 2012a). Floodplain reconnection increases riparian groundwater residence time, which can result in a measurable decrease in in-stream nitrate concentrations (Kaushal et al., 2008). Enhancing lateral connectivity and increasing hydraulic retention time with in-stream and floodplain wetlands can result in significant nitrogen load retention (Browning, 2008; Filoso and Palmer, 2011; Sivirichi et al., 2011). A recent comprehensive review suggested that restoring hydrologic connectivity (e.g. lowering floodplain, raising stream bottom, increasing sinuosity) has generally positive outcomes for nutrient processing and retention (42% - 83% success rate, depending on practice) (Newcomer Johnson et al., 2016).

5.2.2. Nitrogen and vertical connectivity

Although lateral connectivity is an important and visible linkage, vertical connectivity between the stream and subsurface can also influence nutrient retention and removal. In-channel structures (e.g. riffles, cross-vanes, weirs, etc.) can increase nutrient uptake by enhancing in-stream transient storage (Ensign and Doyle, 2005; Baker et al., 2012) and hyporheic exchange (Kasahara and Hill, 2006b;
Hester and Doyle, 2008; Crispell and Endreny, 2009; Azinheira et al., 2014). Hyporheic exchange and subsurface nitrate removal may be significant at constructed riffles (Kasahara and Hill, 2006a, b; Knust and Warwick, 2009), cross-vanes (Daniluk et al., 2013), and other in-channel structures (Hester and Doyle, 2008). In addition, these installed structures can create secondary geomorphic features (e.g. pools and riffles), which can induce greater hyporheic exchange than the structures themselves (Fanelli and Lautz, 2008; Gordon et al., 2013). Direct modification of subsurface hydraulic conductivity has also been suggested as a potential strategy to enhance hyporheic exchange (Vaux, 1968; Hester and Gooseff, 2010; Hester and Cranmer, 2014). Modeling suggests that these types of features would function similarly to natural geomorphic features in driving subsurface flux (Ward et al., 2011).

Alternatively, more natural restoration techniques can also enhance in-stream nutrient removal. In-stream wood can reduce flow velocities (Elosegi et al., 2016), enhance surface transient storage, and provide similar nitrate uptake benefits as more construction-intensive structures (Mueller Price et al., 2016). Installed or natural log jams trap organic matter (Wallace et al., 1995; Quinn et al., 2007) and these organic debris dams have higher denitrification potential compared to other geomorphic features such as gravel bars, pools, and riffles (Groffman et al., 2005; Harrison et al., 2012). Since in-stream wood abundance has often been significantly reduced in urban and degraded streams (Booth et al., 1996; Segura and Booth, 2010), riparian planting may be an appropriate restoration strategy to restore sustainable organic carbon fluxes to streams (Stanley et al., 2012). However, in-stream wood may not be retained in flashy urban streams (Larson et al., 2001), and can be hazardous to infrastructure and boaters.

Beaver reintroduction has been proposed as a viable and cost-effective stream restoration technique (Roni and Beechie, 2013; Pollock et al., 2014). Beaver dams can achieve many restoration objectives including increase geomorphic complexity, serve as grade control, and increase nutrient retention, and have the added benefit of being more resilient than engineered structures. Beaver dams can induce significant hyporheic flux (Fanelli and Lautz, 2008) and beaver ponds can be significant nutrient sinks (Law et al., 2016), especially at high flows (Maret et al., 1987). Furthermore, denitrification in beaver ponds is significant, and they may be important watershed-scale nitrate sinks (Lazar et al., 2015).

On the other hand, removal of anthropogenic dams is an increasingly common restoration practice for increasing longitudinal connectivity and restoring natural flow regimes; however, this may have unintended consequences for nutrient dynamics. Dams create conditions conducive to nutrient retention (e.g. deposition and biotic processing in anoxic bed sediments) and dam removal can decrease nutrient uptake and increase nutrient concentrations (Doyle et al., 2003). Dams may also be important watershed-scale nitrogen sinks. Luckily, many nitrogen-retaining dams are on smaller order streams and removal of dams on larger rivers
may have fewer detrimental effects on nitrogen storage, and greater benefits for fish passage (Gold et al., 2016).

Despite the promise of direct in-channel modifications for increasing hydraulic residence time and hyporheic exchange, significant limitations exist. Only a small fraction (<1%) of the total stream discharge may be exchanged at a given structure (Gordon et al., 2013; Azinheira et al., 2014) and subsurface hydraulic residence times and organic carbon availability may be insufficient for adequate nutrient removal to occur (Gordon et al., 2013; Smidt et al., 2014). Furthermore, bed compaction and use of large footer boulders for structure installation can result in significantly lower hyporheic exchange compared with reference reaches (Becker et al., 2013). Similar limitations exist for floodplains and other forms of lateral connectivity. Nitrate removal in floodplains may be limited due to infrequent inundation (Azinheira et al., 2014; Jones et al., 2015). Furthermore, increased hydrologic connectivity may be insufficient for nutrient retention in areas with very high nitrogen loads (Roley et al., 2012b; Davis et al., 2015; Mahl et al., 2015). Due to small removal rates in the hyporheic zone and floodplains, unreasonably long reaches of stream need to be restored to meet nutrient removal goals (Hester et al., 2016). Finally, floodplains, riparian areas, and the hyporheic zone can all be net sources of nutrients, and more toxic compounds (e.g. mercury), depending on oxygen availability and redox conditions (Vidon et al., 2010; Jones et al., 2015).

5.2.3. Factors controlling denitrification rates

Stream denitrification rates have been shown to be influenced by stream nitrate concentrations (Piña-Ochoa and Álvarez-Cobelas, 2006; Seitzinger et al., 2006; Böhlke et al., 2009; Mulholland et al., 2009; Roley et al., 2012a), land use (Mulholland et al., 2008, 2009), and measurement method (Groffman et al., 2006; Böhlke et al., 2009). Our results confirm the importance of these variables, but ours is the first study that quantifies the relative magnitude of their effects. We found that nitrate is by far the most influential one, while land use and method are of similar importance (Fig. 7). This aligns with previous conclusions that higher denitrification rates in disturbed systems are primarily due to higher nitrate loading (Mulholland et al., 2008).

Climate emerges as the most important variable for riparian denitrification rates, but is only moderately important for streams. Using latitude as a coarse estimate of climate, Piña-Ochoa and Álvarez-Cobelas (2006) found no correlation with stream denitrification. For European riparian sites, denitrification has also been shown to be invariant with climate (Hefting et al., 2004), while others demonstrated some correlation with temperature (Pinay et al., 2007). The difference in the influence of climate on denitrification rates could be due to fundamental process differences between streams and riparian sites; however, other factors may also have influenced the results. For example, the vast majority of the data are from Cf- and Df-climates, temperate, and continental climates with no dry seasons. Additionally, stream sites encompass nine climate types compared to five for riparian sites, which could help explain the relative importance of climate between these two data sets.
The influence of vegetation type on nutrient removal in riparian buffers has been subject to some debate. Our analysis suggests that bare riparian sites tended to have higher denitrification rates than vegetated sites, although the small size of the bare subset (n = 3) makes the significance of this difference suspect. Although riparian vegetation type overall was somewhat influential, it had less of an effect than climate, method, or temporal scale. Others have suggested that total nitrate removal (Groffman and Crawford, 2003; Schoonover and Williard, 2003; Mayer et al., 2007; King et al., 2016) and denitrification rates (Hefting et al., 2004) are not significantly different in buffers with different vegetation types. Similarly, we found no significant difference between denitrification rates in forested and grassed buffers.

Carbon limitation may be a significant factor influencing denitrification in degraded streams, especially those with high nitrogen loading (Lefebvre et al., 2004; Johnson et al., 2009; Newcomer et al., 2012). Although only a limited number of studies in our database examined what was limiting denitrification (carbon, nitrate, or both), there appears to be some correlation with land use. Impacted sites (urban and/or agriculture) had stream denitrification rates that were more often limited by carbon or carbon and nitrate than reference sites (Table 4; chi-squared goodness-of-fit test, \( p = 0.035 \)). This is a tentative conclusion using a limited data set, but the importance of carbon limitation is supported by empirical studies. Carbon deficits in historically disconnected floodplains can lead to lower denitrification potential than hydrologically connected floodplains (Theriot et al., 2013), and experimental carbon addition can increase denitrification rates to levels observed in reference sites (Ullah and Faulkner, 2006). Restored stream reaches may also have lower denitrification rates due to lower organic carbon availability (Gift et al., 2010).

Short-term organic carbon additions, whether experimental or natural (i.e. leaf litter), can enhance removal of both nitrate and phosphorus (Mulholland et al., 1985; Bernhardt and Likens, 2002; Argerich et al., 2008; Aldridge et al., 2009) and enhance subsurface denitrification (Robertson and Merkley, 2009; Zarnetske et al., 2011). More labile forms of carbon (e.g. leaf litter; Aldridge et al., 2010) are likely to be a more important control on in-stream nutrient removal than coarser in-stream wood (Wallace et al., 1995). However, while in-stream wood may not be a significant source of bioavailable carbon, it is effective at trapping other sources of more labile organic carbon (Wallace et al., 1995; Quinn et al., 2007).

**Table 4.** Contingency table of the number of studies with various limiting factors for in-stream denitrification rates by watershed land use type.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Limiting factor</th>
<th>Total</th>
<th>Carbon</th>
<th>Nitrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impacted (Ag or Urban)</td>
<td>Both</td>
<td>9</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Reference</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>
5.3. Implications for stream restoration

The number of studies with sufficient data to confidently assess the performance of stream restoration projects is increasing, but there is still comparatively little information for quantitatively evaluating the relative efficacy of different restoration methods for achieving nutrient reduction goals. The complexity of stream ecosystems and the significant variability among regions, watersheds, and sites further complicates comparisons. Because of these issues, previous reviews of stream restoration used self-reported success of individual studies, rather than some objective measure of restoration benefits. Despite these concerns, Palmer et al. (2014) found that riparian restoration had the highest success in increasing nutrient uptake rates and reducing fluxes (88%), followed by in-stream structure installation (63%), wetland creation (25%), and channel reconstruction (14%). However, sample sizes were small, and reported improvements were not necessarily statistically significant. Newcomer Johnson et al. (2016) found that positive nutrient retention results were reported in 42–83% of projects for restoration strategies that increased hydrologic connectivity (e.g. floodplain connection, increased sinuosity).

While some success has been observed, it is important to recognize the limitations of stream restoration. Degraded streams may have high nutrient uptake rates, comparable to or greater than reference systems (Bernhardt and Palmer, 2007; Mulholland et al., 2008; Johnson et al., 2009). However, this is largely the result of higher nutrient loading to these streams, and removal efficiency is often lower than in undisturbed streams (Mulholland et al., 2008), a conclusion supported by our compiled denitrification data. This has important implications for stream restoration in these disturbed streams. There is likely some maximum nutrient assimilative capacity of streams, and at a certain point, stream restoration will be ineffective in improving water quality. It may be tempting to assume that degraded streams exhibit no nutrient removal and that any stream restoration will be an improvement, but this is unlikely. This is not to say that restoration of degraded streams is not beneficial, only that it is important to consider site-specific context in designing a restoration strategy. For example, a degraded stream with reduced geomorphic complexity may have significant denitrification potential but few active denitrification sites. Increasing the area of sites conducive to denitrification (e.g. hyporheic exchange sites, debris dams, bedforms) can increase total nitrate removal.

The potential constraint on in-stream nutrient processing has led to the suggestion that stream restoration alone may be insufficient for reducing nutrient concentrations to desired levels if no upland best management practices are incorporated (Walsh et al., 2005; Bernhardt and Palmer, 2007; Craig et al., 2008; Selvakumar et al., 2010). Properly designed stormwater controls and source reduction strategies (e.g. fixing leaky sewers, Pennino et al., 2016) can reduce nutrient delivery to streams and potentially improve the success of in-channel restoration efforts (Carboneau and Resh, 1992). There is an increasing recognition that urban stream
restoration requires a watershed-scale approach, addressing causes of degradation (e.g. altered water and sediment regimes) rather than applying the band-aids of reach-scale restoration projects (Roni and Beechie, 2013; Vietz et al., 2016).

There is evidence for nutrient removal in restored streams at small spatial and short temporal scales, but it is uncertain how these observations translate into broader water quality benefits. Removal is limited by inadequate hyporheic exchange at individual in-stream structures, infrequent floodplain inundation, and groundwater bypassing or undercutting riparian zones. Stream restoration projects are generally small scale, and may not result in measurable in-stream benefits (Sutton et al., 2010; Collins et al., 2013; Muller et al., 2016). These results lead to a disconcerting conclusion. While stream restoration can successfully remove nutrients from the stream, the benefits are generally small relative to the scale of the problem, and unreasonably large segments of streams would have to be modified to achieve desirable benefits from restoration alone (Hester et al., 2016).

6. Conclusions and recommendations

Despite many limitations of existing data on the effects of stream restoration on nutrient dynamics, we can make some general conclusions and recommendations:

- Bank stabilization has the greatest potential to reduce phosphorus loading in highly unstable streams, but other restoration strategies have a much smaller effect on phosphorus fluxes.
- Structures that are already incorporated into stream restoration designs can be constructed to encourage hyporheic exchange and nutrient processing through enhanced geomorphic complexity and bedforms, especially during low flow. However, low surface-subsurface exchange rates limit removal potential. The largest nitrogen removal potential likely comes from large-scale manipulations to increase hydraulic retention time (e.g. in-channel and floodplain wetlands and stormwater ponds).
- Riparian buffers remove groundwater nitrate, protect streambanks, supply in-stream wood and organic carbon to increase in-stream processing, and are generally less intrusive and more cost-effective than construction-heavy restoration techniques. However, it is still unclear how significantly buffers affect stream water quality.
- Managing flows of water and sediment and reducing nutrient loading through source controls remains an essential strategy for improving water quality. A combination of restoration strategies that reduce nutrient inputs to streams, re-establish riparian functions, provide balanced water and sediment regimes, and increase in-stream nutrient processing and retention will likely be most effective for improving water quality.

The increase in published stream restoration monitoring studies is encouraging, but the lack of consistent methodology and failure to measure the right metrics makes it difficult to draw conclusions about restoration success (Rubin et al., 2017).
Measuring the benefits of restoration projects is also complicated by the variable nature of in-stream nutrient concentrations; it is difficult to separate the effects of restoration projects from the noise of water quality data. These shortcomings can be addressed by more formal and statistically rigorous monitoring designs. Restoration monitoring should focus on metrics that are specific to project objectives (Palmer et al., 2005) and should follow accepted statistical designs (e.g. Before-After-Control-Impact), comparing restored and control (both impaired and reference) sites over years to decades. Long-term monitoring is required since there may be significant lag time between completion of restoration projects and observed improvements (Meals et al., 2010). Furthermore, monitoring should be well planned and hypothesis driven (e.g. bank stabilization will reduce phosphorus loads by 10%), rather than an afterthought. The complexity of stream restoration lends itself to adaptive management. This is an active approach to resource management, requiring constant monitoring to test the effects of management interventions (e.g. restoration) and updating management approaches as new data are collected (Walters, 1986; Allen et al., 2011). Designing restoration projects as experiments and rigorously monitoring them would reduce uncertainty and increase our knowledge of the effects of restoration on nutrient dynamics.

Existing studies suggest that certain restoration strategies may enhance nutrient retention and removal in stream and riparian systems; however, this is still an emerging science with insufficient evidence to precisely quantify the effects of different restoration strategies across regions and various hydrologic, geologic, and land use contexts. There is an increasing emphasis on improved monitoring of stream restoration success and a growing number of studies on nutrient dynamics in restored systems. Given these trends, the future looks promising for improving our understanding of the role of stream restoration in nutrient management.

Acknowledgments

The authors are grateful to R.W. McDowell, R.C. Walter, D.J. Merritts, G.A. Fox, R.B. Miller, K.E. Schilling, J.L. Kovar, E.O. Young, and D.S. Ross for sharing their streambank phosphorus data. They also thank J. Clary, S. Struck, E. Strecker, A. McGuire, and E. Hester for helpful feedback on a previous version of this manuscript. F.D. Shields and an anonymous reviewer provided comments that greatly improved the quality of this publication.

Funding

This work was partially funded by the Water Environment & Reuse Foundation (Project #WERF-U5R14) and the National Science Foundation, Integrative Graduate Education and Research Traineeship (IGERT) Grant No. DGE-0966346 “I-WATER: Integrated Water, Atmosphere, Ecosystems Education and Research Program” at Colorado State University. Additional funding was provided by the United States Environmental Protection Agency (EPA) Grant 83557001. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the EPA. Further, EPA does not endorse the purchase of any commercial products or services mentioned in the publication.
References


