ABSTRACT: Two-stage ditches represent an emerging management strategy in artificially drained agricultural landscapes that mimics natural floodplains and has the potential to improve water quality. We assessed the potential for the two-stage ditch to reduce sediment and nutrient export by measuring water column turbidity, nitrate (NO$_3^-$), ammonium (NH$_4^+$), and soluble reactive phosphorus (SRP) concentrations, and denitrification rates. During 2009-2010, we compared reaches with two-stage floodplains to upstream reaches with conventional trapezoid design in six agricultural streams. At base flow, these short two-stage reaches (<600 m) reduced SRP concentrations by 3-53%, but did not significantly reduce NO$_3^-$ concentrations due to very high NO$_3^-$ loads. The two-stage also decreased turbidity by 15-82%, suggesting reduced suspended sediment export during floodplain inundation. Reach-scale N-removal increased 3-24 fold during inundation due to increased bioreactive surface area with high floodplain denitrification rates. Inundation frequency varied with bench height, with lower benches being flooded more frequently, resulting in higher annual N-removal. We also found both soil organic matter and denitrification rates were higher on older floodplains. Finally, influence of the two-stage varied among streams and years due to variation in stream discharge, nutrient loads, and denitrification rates, which should be considered during implementation to optimize potential water quality benefits.

(KEY TERMS: best management practices; two-stage ditch; restoration; biogeochemistry; denitrification; nutrients; turbidity; agricultural stream.)

INTRODUCTION

More than one-third of rivers in the United States (U.S.) (USEPA, 2000) and two-thirds of the world’s estuaries (Bricker et al., 2007) are significantly impaired or polluted, largely due to excess nitrogen (N) and phosphorus (P) from anthropogenic sources. Biologically available N inputs to the biosphere have doubled since the advent of synthetic fertilizer production (Vitousek et al., 1997), and have been linked to locally contaminated drinking water (Fan and Steinberg, 1996), loss of freshwater biodiversity (Carpenter et al., 1998), and coastal eutrophication and hypoxia (Rabalais et al., 2002). In addition, anthropogenic P additions have been shown to drive eutrophication in inland freshwaters (Schindler, 1977).
The majority of nutrient export from the U.S. Midwest comes from small streams draining agricultural fields (Carpenter et al., 1998), which contribute an estimated 70% of the annual N load to the Gulf of Mexico (Alexander et al., 2008). This excess nutrient loading is associated with high N fertilizer application rates and the predominance of artificial drainage (David et al., 1997; USGS, 1999). Surface drainage in the Midwestern agricultural landscape is conventionally maintained through regular dredging of streams and ditches, which results in incised channels that tend to be unstable during high flows (Figure 1A) (Powell et al., 2007b). Additionally, >20 million hectares of agricultural soils in the Mississippi River Basin are drained by subsurface tile drains which discharge directly into these channelized streams and ditches (Osborne and Wiley, 1988; Sugg, 2007). The combination of channelization and tile drainage improves crop production by rapidly conveying excess water from fields (Fausey et al., 1995), but also reduces water contact time with soils and sediments, which ultimately reduces retention of excess N and P (Randall et al., 1997; Royer et al., 2006).

The two-stage ditch is an alternative management strategy that can be applied to conventional channelized ditches in order to maintain drainage while simultaneously reducing nutrient export. In the two-stage ditch, floodplain benches are constructed alongside the main stream channel by excavating a portion of the riparian zone, often used for the conventional grass buffer strip (Figures 1A and 1B) (Powell et al., 2007b). The two-stage design mimics the natural floodplain benches that form in unmaintained ditches as the result of common fluvial processes (e.g., bank slumping, undercutting, and sediment deposition; Landwehr and Rhoads, 2003). During storms, inundated floodplains, be they natural or constructed, allow water to spread out and slow down, which reduces shear stress and results in improved channel stability and reduced erosion (Powell et al., 2007a). In addition to improvements in channel stability, constructed floodplains also have the potential to reduce nutrient export by increasing water residence time during floodplain inundation, which allows for increased sediment deposition, enhanced assimilation of N and P by vegetation, and increased rates of per-
permanent N-removal through microbial denitrification (Fennessy and Cronk, 1997). In addition, two-stage floodplains could also retain more nutrients than conventional riparian buffers at base flow because floodplains are closer to the water table and are hydrologically connected to surface water (Hill, 1996; Groffman and Crawford, 2003; Roley et al., 2012b). Furthermore, subsurface flow shunted through tile drains generally bypasses upland buffer strips, and can be intercepted by two-stage floodplains before entering the main channel (Osborne and Kovacic, 1993).

Research evaluating the efficacy of the two-stage ditch practice has largely focused on fluvial processes and channel stability (Powell et al., 2006, 2007a, b), while research related to potential water quality benefits of two-stage ditches has been limited to a few recently constructed sites (Roley et al., 2012a, b; Davis et al., 2015). Nutrient retention capacity of two-stage floodplains depends on the interaction between varying process rates and water residence times. For example, rates of microbially mediated denitrification are primarily limited by substrate availability (i.e., denitrification requires nitrate [NO$_3^-$] and carbon) and secondarily by physical factors (e.g., hydrology, temperature, redox conditions). Therefore, we predict that the effects of two-stage construction will vary across sites and over time due to variation in nutrient loads and discharge-mediated inundation frequency. Function of two-stage floodplains should also improve as floodplain vegetation matures and organic matter accumulates over time. Finally, we predict that reach-scale effects of two-stage construction will depend on flood frequency and water residence times, which are controlled by floodplain bench height and surface area.

Our objectives were to determine how the potential water quality benefits of the two-stage practice are influenced by: (1) physiochemical conditions across sites (i.e., nutrient loading and discharge), (2) dimensions of constructed floodplains, and (3) evolution of the two-stage ditch over time using six two-stage ditches constructed in Midwestern streams that varied in age from 1 to 10 years post-construction. Studies that evaluate restorations across sites and through time are rare but have the potential to provide critical insight for future site selection and restoration design (Moerke and Lambert, 2004; Palmer et al., 2005). Results from this study indicate that the effectiveness of the two-stage ditch depends on the restoration goal (i.e., substrate improvement vs. sediment, P, or N retention). We identify environmental constraints on denitrification rates and demonstrate the potential to increase N-removal at the reach scale by constructing floodplains at a low height, which can increase inundation duration without compromising channel stability.

METHODS

Site Description

We chose six streams with two-stage reaches located within agriculturally dominated watersheds (>70% of watershed area in agriculture) in Indiana, Michigan, and Ohio, U.S. (Figures 1B, 1C, and 2). At all sites, adjacent fields are tile-drained and planted in row crops, primarily as a corn-soybean rotation, and the streams have historically been maintained as drainage ditches with steep slopes and incised, trapezoidal channels (Figure 1A). As a result, the study streams have flashy hydrographs (e.g., Figure 1, Griffiths et al., 2013) and high concentrations of dissolved inorganic nutrients. Two-stage floodplains were constructed from 1 to 7 years prior to our study. For a baseline, we also included a stream that had not been maintained (i.e., no routine channelization) for >10 years prior to our study; thus this system had floodplain benches that had formed naturally. Together, our study sites capture a range of nutrient loads and floodplain designs (i.e., varying width and height; Table 1, Figure 2). In addition, we paired each two-stage reach (hereafter referred to as two-stage) with an upstream control reach (hereafter referred to as reference), representing a typical incised, trapezoidal channel without floodplains, and throughout the study we sampled both the reference and two-stage reaches (Figure 1D). We measured denitrification rates seasonally over one year, between summer 2009 and spring 2010. We collected surface water samples more frequently, based on site visits, and reported means based on all samples collected in 2009-2010 (Table 2).

Stream Benthic Habitat

We used a modified Wolman pebble count to quantify stream bottom habitat during the summer in both 2009 and 2010. We classified sediments from 20 transects that were located at evenly spaced intervals along each stream reach (i.e., upstream reference reach and downstream two-stage reach), and categorized substrate as clay, fine benthic organic matter (FBOM), sand (0.01-2 mm), gravel (2-22 mm), or cobble (>22 mm) (sensu Cummins, 1962). For each transect, we sampled sediments at 10-cm intervals across the stream width (i.e., perpendicular to flow).
Streamflow and Stage

We deployed capacitance meters (Odyssey, Christchurch New Zealand), which logged stream stage at 10-30 min intervals, at the downstream end of the reference and two-stage reaches of each stream and also on the floodplain benches in the two-stage reaches. We then correlated reach-specific capacitance data with stream discharge data from USGS gauges near our sites exhibiting hydrographs that best matched site conditions, and used a watershed yield method to estimate site discharge ($Q_{site}$) based on relative drainage areas (DA) for the site and the gauge using the equation $Q_{site} = Q_{gauge} \times (\frac{DA_{site}}{DA_{gauge}})$ (Roley et al., 2014). We estimated drainage areas for sites in Indiana and Ohio using the web-based StreamStats tool (http://water.usgs.gov/osw/streamstats/ssonline.html) and the Digital Watersheds tool (http://www.iwr.msu.edu/dw/) for the Michigan site (CRO). We determined floodplain inundation stage based on data from capacitance meters located on floodplain benches and confirmed these data based on field observations (i.e., stage when streams were at bankfull, floodplains were inundated, or debris dams provided evidence of recent inundation).

Stream Turbidity and Nutrient Concentrations

We deployed Hydrolab MS5 Minisondes (Hach Company, Loveland, Colorado) at the downstream ends of the two-stage and upstream reference reach in each stream. Sondes were deployed in perforated PVC tubes and were suspended above the stream bottom parallel to the main channel flow to prevent clogging due to sedimentation. We calibrated sondes approximately every 2 weeks, and they recorded dissolved oxygen (DO in mg/L), temperature (°C), and turbidity (in NTUs) every 30 min. All measurements fell well within ranges of detection which were from 0 to 60 mg DO/L, 0-30°C, and 0-3,000 NTUs, respectively.

We collected surface water grab samples for nutrient analysis from the downstream end of each reach whenever we were at sites during 2009 to 2010. We immediately filtered surface water samples into 60 ml HDPE bottles using type AE GFF syringe-mounted filters and stored the samples on ice prior to returning to the laboratory, where samples were frozen at −24°C prior to analysis. We analyzed samples for NO$_3^-$ by cadmium reduction (APHA, 1995), ammonium (NH$_4^+$) by the phenol-hypochlorite method (Solórzano, 1969), and soluble reactive phosphorus (SRP) using the molybdate blue method (Murphy and Riley, 1962), on a Lachat Quickchem 8500 Flow Injection Analyzer (Hach Company). Our detec-
tion limits ranged from 2-2,000 μg NO₃⁻-N/L, 1-200 μg SRP-P/L, and 2-1,000 μg NH₄⁺-N/L, and we diluted samples as needed to optimize analysis.

Field Sampling of Stream Sediments and Floodplain Soils

We collected stream sediments using a PVC corer (3.6 cm diameter) and floodplain soils using an Oakfield corer (1.5 cm diameter) once per season from each site for one year (from spring 2009 through summer 2010). We sampled to a depth of 2 cm for stream sediments and 5 cm for floodplain soils because denitrification rates decrease substantially below these depths in stream sediments (Arango et al., 2008) and in soils on floodplain benches (Roley et al., 2012b). To ensure that samples were representative of site-specific substrate heterogeneity, we collected eight replicate samples at evenly spaced intervals along each reach and, for each replicate, we pooled multiple cores from transects that spanned the width of either the inset stream channel or the floodplain. We stratified floodplain sampling between vegetated and unvegetated areas for recently constructed sites that were not completely vegetated (CRE, KLA, CRO). Sediment and soil samples were placed on ice and returned to the laboratory for denitrification assays.

Laboratory Assays for Denitrification

Within 24 h of sediment and soil collection, we measured microbial denitrification rates using the redox-optimized chloramphenicol-amended acetylene block technique (Smith and Tiedje, 1979; Royer et al., 2004; Arango et al., 2007). Acetylene prevents the conversion of nitrous oxide (N₂O) to dinitrogen gas (N₂), resulting in N₂O accumulation in assay bottles, which is measurable on a gas chromatograph. Because the addition of chloramphenicol prevents de novo production of denitrifying enzymes in response to optimized redox conditions (Brock, 1961), the addition of chloramphenicol minimizes bottle effects associated with laboratory assays, particularly for shorter incubation periods (Bernot et al., 2003).

We homogenized sediment samples and cut floodplain soil samples into small pieces because this facilitates C₂H₂ diffusion when soils are too moist for sieving (Groffman et al., 1999). We then added 25 mL of sediment or soil, 45 mL of unfiltered stream water, and 5 mL of 3.1 mM chloramphenicol to glass media bottles (Bruesewitz et al., 2008; Roye et al., 2012a, b). Next, we sealed bottles with septum caps and sparged with ultra-high purity N₂ gas for 5 min, swirling periodically to remove oxygen. We then vented bottles to return them to atmospheric pressure before adding 15 mL of C₂H₂ gas. We incubated assay bottles at room temperature for a total of 4 hr, taking 5 mL gas samples from the headspace approximately every hour. We shook each bottle prior to sampling to equilibrate gases in the headspace and slurry, then removed 5 mL of gas with a syringe, injected the gas into evacuated 3 mL serum vials with rubber septa (Wheaton, Millville, New Jersey), and replaced the sample volume with 5 mL of 10% C₂H₂ to maintain constant pressure within each assay bottle.

In addition to running assays at ambient chemical conditions (i.e., no added NO₃⁻ or organic carbon), we also wanted to determine whether denitrification rates were limited by NO₃⁻ or organic carbon availability. Therefore, we conducted nutrient limitation

<table>
<thead>
<tr>
<th>Site</th>
<th>State</th>
<th>Watershed</th>
<th>Reach</th>
<th>SRP (μg/L)</th>
<th>Mean ± SE</th>
<th>Range</th>
<th>NH₄⁺-N (μg/L)</th>
<th>Mean ± SE</th>
<th>Range</th>
<th>NO₃⁻-N (μg/L)</th>
<th>Mean ± SE</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHA IN</td>
<td>Tippecanoe</td>
<td>REF</td>
<td></td>
<td>31 ± 6</td>
<td>8-129</td>
<td></td>
<td>50 ± 9</td>
<td>9-212</td>
<td></td>
<td>5,355 ± 373</td>
<td>2,072-10,447</td>
<td>26</td>
</tr>
<tr>
<td>CRE IN</td>
<td>St. Joseph</td>
<td>River</td>
<td>TRT</td>
<td>21 ± 5</td>
<td>8-116</td>
<td></td>
<td>87 ± 8</td>
<td>30-180</td>
<td></td>
<td>5,424 ± 331</td>
<td>1,926-9,170</td>
<td>15</td>
</tr>
<tr>
<td>CRO MI</td>
<td>St. Joseph</td>
<td>River</td>
<td>TRT</td>
<td>28 ± 7</td>
<td>10-129</td>
<td></td>
<td>167 ± 46</td>
<td>6-710</td>
<td></td>
<td>1,145 ± 303</td>
<td>3-4,563</td>
<td>18</td>
</tr>
<tr>
<td>KLA OH</td>
<td>Portage</td>
<td>TRT</td>
<td></td>
<td>61 ± 28</td>
<td>6-325</td>
<td></td>
<td>212 ± 100</td>
<td>7-1,826</td>
<td></td>
<td>1,121 ± 304</td>
<td>5-4,560</td>
<td>18</td>
</tr>
<tr>
<td>BUL OH</td>
<td>Upper Miami</td>
<td>TRT</td>
<td></td>
<td>58 ± 28</td>
<td>8-322</td>
<td></td>
<td>73 ± 12</td>
<td>22-161</td>
<td></td>
<td>6,669 ± 1,313</td>
<td>39-13,906</td>
<td>11</td>
</tr>
<tr>
<td>CEAP OH</td>
<td>Upper Scioto</td>
<td>TRT</td>
<td></td>
<td>60 ± 24</td>
<td>13-148</td>
<td></td>
<td>70 ± 32</td>
<td>27-196</td>
<td></td>
<td>2,065 ± 535</td>
<td>921-3,610</td>
<td>6</td>
</tr>
</tbody>
</table>

Note: For each site and reach combination, the nutrient concentration is reported as Mean ± Standard Error (SE) for all samples taken during 2009-2010. Samples were collected primarily at base flow and analyzed for dissolved nutrients (SRP, NH₄⁺, NO₃⁻). Duration of sampling varied between sites but included at least one day per season for all sites from summer 2009-spring 2010; n = replicate number.
assays where we added KNO₃⁻ or organic carbon (as glucose) alone and in combination to create three additional denitrification assay treatments: +N, +C, +N+C. We prepared assays as described above but increased NO₃⁻ concentrations by 1.1 mg NO₃⁻-N/L above background (for +N and +N+C) and carbon by 2.7 mg C/L above background (Bruesewitz et al., 2008; Roley et al., 2012a, b).

We analyzed gas samples collected for all denitrification assays using a Varian CP-3800 gas chromatography.
Two-Stage Ditch Floodplains Enhance N-Removal Capacity and Reduce Turbidity and Dissolved P in Agricultural Streams

Statistical Analyses

To determine the influence of the two-stage on surface water nutrients, we performed one-tailed paired t-tests that compared the upstream reference and downstream two-stage reaches at each site, paired by date. When differences based on all site data were significant, we also did paired t-tests for individual sites. If data were non-normal, we transformed the data or used the nonparametric paired Wilcoxon test if normality could not be achieved.

We compared denitrification rates for upstream reference and downstream two-stage reaches of the main channel and two-stage floodplains using a nested analysis of variance (ANOVA) for which we blocked each reach within the site. We used simple linear regression to test relationships between denitrification rates and physiochemical variables (e.g., NO$_3^-$ and soil or sediment organic matter content). To determine whether denitrification was nutrient-limited, we used a two-way ANOVA on data from denitrification nutrient limitation assays (Tank and Dodds, 2003). We transformed data as needed to meet assumptions for normality and equal variance, using the log or square root of the data or rank-sum transformation (Iman and Conover, 1979). We performed statistical analysis using SYSTAT 12 (SYSTAT Software, 2007) or R version 3.0 (R Core Development Team, 2012), and considered $p < 0.05$ to be significant, reporting most values as the mean ± standard error (SE).

RESULTS

Variation in Streamflow and Floodplain Inundation Patterns

During the study period from summer 2009 through spring 2010, mean base flow for the sites in Indiana and Michigan (SHA, CRE, and CRO) ranged from <30 L/s in Fall 2009 to >100 L/s in spring 2010, while peak flow was ~2,000 L/s (Table 1). Sites in Ohio (KLA, BUL, and CEAP) had lower base flow (0.2-44 L/s) and peak flows (~1,000 L/s) (Table 1). We note that in general, discharge during 2009 was below average for sites in Ohio, while 2010 was considered a drought year for all sites, based on long-term USGS flow gauge data for our watersheds (http://waterdata.usgs.gov).

Width of constructed floodplains was based on the minimum width required for the ditch to contain peak flows (i.e., 200 year flood; see Powell et al., 2007a, b) and generally increased stream area by three to five fold during floodplain inundation. Mean depth of the main channel (i.e., floodplain height) ranged from 0.21 m to 0.43 m across all sites (Table 1). Bankfull discharge, above which the stream water inundates two-stage floodplains, ranged from 65 to 300 L/s. This threshold discharge increased with bench height, resulting in fewer inundated days each year in streams with higher benches (Figures 3A and 3B). Mean inundation event frequency during the study period ranged from 9 to 14 events/year but event duration was highly variable (Table 1), with event frequency and duration being, especially high in spring and winter. From 2009 to 2010, conditions were relatively dry and annual duration of inundation ranged from 10 to 116 day/yr, depending on bench height and drainage area (Table 1). Wetter conditions with more storms resulted in twice as many flood days in 2011, and the lowest floodplains (i.e., KLA, BUL, and SHA) were inundated for more than a third of the year (Figure 3B). The CEAP site, where floodplains were allowed to form naturally due to a lack of conventional channel maintenance, had one of the highest benches, the smallest drainage area, and was inundated least frequently.

Graph (Varian, Walnut Creek, California) that was equipped with an electron capture detector (ECD), a Haye SepQ column (AllTech, Deerfield, Illinois), a valve to vent water and C$_2$H$_2$ away from the detector, and a CombiPAL autosampler (CTC Analytics; Zwingen, Switzerland). We measured N$_2$O with the ECD, using ultra-high purity N$_2$ as the carrier gas, with the injector and oven set at 50°C, and the ECD set at 300°C.

We calculated total concentration of N$_2$O at each sampling period using the appropriate Bunsen coefficient to determine the amount of gas dissolved in water at a given headspace concentration (Tiedje, 1982; Tiedje and Groffman, 1989; Inwood et al., 2005), then plotted N$_2$O mass vs. time and used the slope of the best-fit linear regression line to determine production rates for denitrification. We scaled production rates from µg N$_2$O-N/h to production per g dry mass (DM), ash free dry mass (AFDM), and area (m$^2$). We measured soil wet mass prior to assays and measured DM (dried at 60°C) and AFDM (measured by loss on ignition) after completing assays (Steinman et al., 2006). We also estimated potential N removal at the reach scale (g NO$_3^-$-N/km/day) during floodplain inundation by multiplying average areal denitrification rates (g N$_2$O-N/m$^2$/day) for the main channel and two-stage floodplain by their respective areas at each site (m$^2$ per 1 km reach; Roley et al., 2012b).
Two-Stage Effects on Benthic Substrate Composition in the Main Channel

Due to the two-stage design, stream velocities should be higher in the main channel after floodplain construction, which could influence substrate composition and the accumulation of fine sediments. In general, the main inset stream channels in the older (>4 years) two-stage ditches had more (8-30%) coarse substrate when compared to upstream, reference reaches, but we found minimal or no increase in coarse substrate at more recently constructed sites (Figure 4). The coarser substrate consisted primarily of gravel (diameter = 2-22 mm), and increases in cobble only occurred at the oldest sites. Conversely, recently constructed two-stages had more FBOM while older two-stages generally had less FBOM compared to reference reaches (Figure 4).

Water Quality across Sites and Effect of the Two-Stage

Concentrations of SRP, NH$_4^+$, and NO$_3^-$ varied both seasonally and among sites, ranging from near minimum detection limits to peaks exceeding 325 µg SRP/L, 1,800 µg NH$_4^+$-N/L, and 13,000 µg NO$_3^-$-N/L (Table 2). Nutrient concentrations were generally highest at sites in the Tippecanoe and Portage watersheds (SHA and KLA), where mean nutrient concentrations were around 30-50 µg SRP/L, 50-100 µg NH$_4^+$-N/L, and greater than 5,000 µg NO$_3^-$-N/L (Table 2). Surface water NO$_3^-$ concentrations were lowest at the St. Joseph River sites (CRE, CRO: means = 509-1,121 µg NO$_3^-$-N/L, respectively). Mean NH$_4^+$ concentrations ranged from 22 to 212 µg NH$_4^+$-N/L, with the lowest and highest means being measured at CRE and CRO, respectively. SRP varied less, with means ranging from 11 to 60 µg SRP/L (Table 2).
Overall, concentrations of SRP declined significantly in the two-stage but changes in NH$_4^+$ and NO$_3^-$ were not significant (Figures 5A, 5C, and 5E). On average, concentrations of SRP were 23% lower at the end of the two-stage (26 ± 5 µg SRP-P/L) than for the reference reach (33 ± 5 µg SRP-P/L). However, decreases in SRP varied across sites and were only statistically significant for CRE, SHA, and CEAP (paired Wilcoxon test, $p = 0.005, 0.01$, and $0.03$, respectively; Table 2, Figure 5A). We note that in KLA, the site with the lowest benches, NO$_3^-$ concentrations at the bottom of the two-stage reduced by 4%, despite very high loading (paired $t$-test, $p = 0.035$; Table 2, Figure 5C). When concentrations were converted to nutrient flux, using instantaneous discharge estimates for the two-stage and upstream reference reaches, we found no significant reductions in nutrient flux for any solute as a result of two-stage construction (Figures 5B, 5D, and 5F).

Finally, we examined whether water column turbidity, as an indicator of suspended sediment load, would be influenced by two-stage construction. During periods of floodplain inundation, the two-stage reach significantly reduced water column turbidity relative to upstream channelized reaches for all streams except for CRE, which had the most recent two-stage construction (Figure 6, paired $t$-test, $p < 0.05$). In contrast, water column turbidity did not differ significantly between reaches at base flow, when floodplains were dry.
Microbial Denitrification Rates in Main Channel vs. Floodplain Benches

Denitrification rates measured on sediments from the main stream channel ranged from 0 to 1.7 μg N₂O-N/g DM/h, with the lowest rates during summer and fall and the highest rates in spring when NO₃⁻-availability was also highest (Figures 7A and 7B). While there were significant differences across seasons and among sites (ANOVA, p < 0.001), stream sediment denitrification rates were consistently higher in two-stage reaches (0.42 ± 0.04 μg N₂O-N/g DM/h) than in upstream reference reaches (0.25 ± 0.03 μg N₂O-N/g DM/h; ANOVA, p = 0.003; Figure 7A). Variation in sediment denitrification rates (per g DM) was explained by stream NO₃⁻ concentrations (Figures 8A and 8B) (SLR, square root transformed: r² = 0.82, p < 0.001) but there was no relationship between denitrification and sediment OM content. Nutrient limitation assays showed similar results, which indicated that sediment denitrification rates were N-limited (Figure 9A). As such, when stream NO₃⁻ concentrations were <5 mg NO₃⁻-N/L, an increase of +1 mg NO₃⁻-N/L resulted in a ~2-24 fold increase in denitrification rates (Figures 9A and 9B). In contrast, when stream NO₃⁻ concentrations >5 mg NO₃⁻-N/L, denitrification rates were not limited by N or C (Figure 9B).

Denitrification rates in floodplain soils ranged from 0 to 3.7 μg N₂O-N/g DM/h and followed seasonal patterns similar to sediments from the main channels (ANOVA, p < 0.001). Floodplain denitrification rates averaged 0.6 ± 0.05 μg N₂O-N/g DM/h and were 1.4 times higher than rates for sediments from the main channel (Figure 7A) (ANOVA, p < 0.001). In contrast to stream sediments, variation in floodplain denitrification rates was predicted by both soil OM (SLR, r² = 0.47, p < 0.001) and NO₃⁻ (SLR, r² = 0.40, p = 0.002), but correlations with NO₃⁻ were not as strong as with stream sediments (Figures 8C and 8D). In general, both soil OM content and denitrification rates tended to be higher in soils that had vegetation compared to unvegetated soils (Figure 8D). Similar to stream sediments, floodplain soil denitrification was N-limited (Figure 9A), but the addition of NO₃⁻ only enhanced denitrification when concentrations were <1 mg NO₃⁻-N/L (Figure 9A). Despite a strong correlation between soil OM and floodplain denitrification rates, nutrient limitation assays indicated that floodplain soils were neither C-limited nor N+C co-limited (Figure 9B).

Scaling Denitrification N Removal up to the Reach

When we expressed denitrification rates per unit area of stream bottom, and compared these to previously published rates, we found that the range of denitrification at our sites (0-53 mg N₂O-N/m²/h) spanned the full range of previously reported values (Figure 10). In contrast, the slope of increase in denitrification in response to increasing stream NO₃⁻ was significantly higher for our Midwestern sites than for other sites across the U.S. with variable land use (i.e., urban, agricultural, and reference sites together; ANCOVA, rank transformed, p = 0.001) (Figure 10). We also estimated reach-scale N-removal using respective reach areas and areal denitrification.
rates for both the main channel and floodplain benches in each stream. We estimated that reach scale N-removal ranged from 1.8 to 10.9 kg N/km/day during floodplain inundation and was 3-24 times higher than N-removal occurring in respective main channels alone (Figure 11C). In addition, reach-scale
N-removal rates and soil organic matter content generally increased with two-stage age (Figure 11A).

**DISCUSSION**

**Two-Stage Ditch Design and Floodplain Inundation**

The efficacy of the two-stage ditch design is mediated by the height and width of constructed floodplains, which will influence their ability to: (1) stabilize banks, (2) sustain drainage capacity, and (3) retain sediment and nutrients. Regional curves, which relate drainage area to channel morphology at sites where floodplain benches have formed naturally, can be used to identify a range of heights at which constructed floodplains should remain stable (Powell et al., 2006). However, nutrient retention by floodplains may be limited by infrequent inundation. For example, annual duration of inundation at our site where floodplains formed naturally (CEAP; Figure 3) was consistent with the range reported for other sites in Ohio (9-26 day/yr) (Powell et al., 2006), whereas annual inundation duration was longer for other two-stage ditch sites (22-141 day/yr; Figure 3) because constructed floodplains were 20-60% lower than the heights recommended for optimal stability based on regional curves (Kallio, 2010). Despite low floodplain height, no additional maintenance was required to sustain drainage and all two-stage floodplains remained stable over time (Kallio, 2010; D’Ambrosio et al., 2015).

For the two-stage to be self-sustaining, it must maintain a dynamic equilibrium between depositional and erosional processes. Agricultural streams are flashy because runoff increases magnitude and frequency of peak flows (Ward and Trimble, 2003), and discharge at our sites increased by more than 10-fold between base and peak flow (Table 1). The number of days that floodplains were inundated also doubled between 2010, which was relatively dry, and 2011, which was the wettest year on record (Figure 3). Optimal dimensions for the two-stage must therefore be based on a range in discharge while potential flood regimes (e.g., inundation frequency and duration)

---

**FIGURE 10**. Relationship between Instream Denitrification Rates and Stream Water NO$_3^-$ Concentrations Measured in This Study Compared with Previously Published Data from Other Systems, Using Seasonal Means (summer, fall, winter, spring) for Each of Our Sites (CRE, SHA, KLA, CRO, BUL, CEAP). Rates from other studies were measured using multiple methods including whole stream 15NO$_3^-$ releases, membrane inlet mass spectrometry (MIMS), in situ chamber techniques, and laboratory C$_2$H$_4$ slurries (see Roley et al., 2012a for literature review). Significant correlations are denoted by * (SLR, $p < 0.05$).

**FIGURE 11.** (A) Areal Denitrification Rates for Floodplain Soils and Soil Organic Matter Content Tended to Increase with Age of the Two-Stage. (B) In contrast, instream denitrification rates tended to vary with stream NO$_3^-$ concentrations. (C) Finally, reach-scale N-removal rates during floodplain inundation were estimated by multiplying denitrification rates for the main channel and floodplains by their respective surface areas (see Table 2). The factor by which floodplains increased total N-removal relative to the main channel are shown numerically above bars. * To facilitate comparison with denitrification rates, means for stream NO$_3^-$ concentration only include data for surface water used in denitrification assays.
could influence channel stability as well as floodplain nutrient retention.

Two-Stage Ditches Had Limited Effects on Stream Substrate

We predicted that the two-stage ditch could improve stream substrate composition because increased velocities in the main channel would increase sediment conveyance and sorting, exposing coarse substrates in the main channel while depositing fines on floodplains (Powell et al., 2007b). Using geomorphic surveys, we found that reductions in FBOM and increases in coarse substrates only occurred at sites with floodplains >5 years old (Figure 4), and thus there is likely a delay in seeing the substrate benefits associated with two-stage construction. The presence of coarse substrates could improve habitat by supporting greater primary production (Uehlinger et al., 2000; Atkinson et al., 2008) and increased diversity and biomass of fish and macroinvertebrates (Grimm and Fischer, 1984; Owens et al., 2005). However, substrates in agricultural streams are typically dominated by finer sediments while the presence of coarse sediments is limited (Fischer et al., 2010; Smiley et al., 2011), which ultimately may limit the potential for habitat improvement.

Two-Stage Ditches Could Potentially Reduce NO₃⁻ and SRP in Surface Water

Agricultural streams generally have elevated nutrient concentrations resulting from excess fertilizer runoff from the surrounding landscape (Royer et al., 2004, 2006) and our study streams were no exception to this trend, especially for NO₃⁻ (Table 2). Given this context, we found that construction of the two-stage ditch over short reaches (i.e., <1 km) did not significantly reduce NO₃⁻ concentrations or loads at most sites (Figures 5E and 5F), despite high potential rates of denitrification (Figures 10 and 11). However, the two-stage site with the lowest benches (KLA) consistently decreased surface water NO₃⁻ concentrations (Figure 5C). In addition to more frequent inundation of floodplains (Figure 3A), tile drain outflow was retained on benches and traveled along extended flow paths (2-10 m) before reaching the stream. Denitrification rates for soils inundated by tile flow were also two times higher than outside flow paths, suggesting that the two-stage could reduce NO₃⁻ loading if two-stage reaches were longer, and floodplain were constructed to retain tile water at base flow.

In contrast to NO₃⁻, we found that relatively short reaches of two-stage could reduce surface water SRP concentrations (Figure 5A), though the magnitude of this reduction was site-specific. For example, mean reductions in SRP concentrations at base flow ranged from 3 to 53% but were significant for only half of the streams, and reductions in SRP flux were not statistically significant (Figures 5A and 5B). Mechanisms for improved P retention could be through abiotic sorption to fine particles (Meyer and Likens, 1979), or through biological uptake where P is assimilated in plant tissues during growth and later released during decomposition. We collected surface water primarily at base flow but expect floodplains to have a greater effect on surface water nutrient concentrations during inundation (Mayer et al., 2007; Noe and Hupp, 2009). Finally, stream substrate composition and vegetation density, which was 17-36% higher in main channels of the two-stage than in reference reaches, could have influenced SRP retention.

Floodplain Inundation Reduced Sediment Export

We also found significant reductions in turbidity (Figure 6), which is a measure of water cloudiness that can correlate with total suspended solids (Jones et al., 2011). These results support our prediction that floodplain inundation reduces stream velocities along margins and results in increased deposition of suspended particles (Woltemade and Potter, 1994). We found that turbidity reductions were highest at the site with the widest floodplains (SHA) and were relatively low (<25%) at sites with narrower floodplains (KLA, BUL; Table 2 and Figure 6). Sediment retention could also improve over time due to establishment of vegetation, which stabilizes soils and further reduces water velocities during inundation (Woltemade and Potter, 1994). For example, at SHA, turbidity reductions increased from 40% in 2008 (S.S. Roley, 2008, unpublished data) to 74% in 2010 (Figure 6). Given that greater than 90% of the sediment load can be exported during high flow events (Withers and Sharpley, 2008; Banner et al., 2009), turbidity reductions would likely drive significant reductions in annual sediment export.

Instream Denitrification Rates Depend on Variation in Surface Water NO₃⁻

Sediment denitrification rates measured at our sites ranged from among the lowest to the highest values previously reported for sites across the U.S.,
and demonstrated potential effects of constructed floodplains on N removal in Midwestern streams with a wide range of NO$_3^-$ concentrations (1-13,000 $\mu$g NO$_3^-$/N/L; Table 2). Consistent with previous research, sediment denitrification rates increased with stream NO$_3^-$ concentrations (Figure 8A; Piña-Ochoa and Álvarez-Cobelas, 2006; Hall et al., 2009). The highest rates of removal therefore coincided with the highest NO$_3^-$ export, which generally occurred during spring, when high flow and storms typically increase runoff from fields (Royer et al., 2006). However, denitrification efficiency (i.e., percent of NO$_3^-$ load removed) can decrease with stream NO$_3^-$ concentration because the ability of denitrifiers to respond to high NO$_3^-$ pulses can become saturated at high concentrations (Inwood et al., 2005). The threshold above which sediment denitrification ceased to be N-limited at our sites was fairly high at 5 mg NO$_3^-$-N/L (Figure 9A), but stream NO$_3^-$ still routinely exceeded this threshold (Table 2). These results suggest that denitrification efficiency will be highest for two-stage floodplains constructed at sites with intermediate NO$_3^-$ loads, and perhaps the two-stage might be optimally paired with on-field management to reduce NO$_3^-$ inputs.

Denitrification can be limited by organic carbon availability when NO$_3^-$ concentrations are high (Arango et al., 2007). Therefore, we predicted that two-stage implementation could reduce denitrification rates in the main channel by reducing standing stocks of organic-rich fine particles (FBOM). We found that FBOM standing stocks were indeed lower at older two-stage sites (>5 years, Figure 4). Nevertheless, and contrary to our predictions, we found that sediment denitrification was consistently higher in the two-stage reaches at all streams except for BUL (Figure 7A) and there was no obvious relationship between denitrification rates and organic matter content for stream sediments (Figure 8B). As Powell and Bouchard (2010) hypothesized in a previous publication, it could be that the absence of ditch maintenance, which would have disturbed vegetation and sediments via dredging, could have promoted undisturbed microbial assemblages and subsequently high denitrification in two-stage reaches.

**Floodplain Denitrification Rates Depend on Accumulation of OM over Time**

In contrast to stream sediments, soil denitrification rates on two-stage floodplains were consistently controlled by soil organic matter content (Figure 8D), which agrees with results from previous studies (Groffman and Crawford, 2003; Orr et al., 2007; Gift et al., 2010). In addition, soil organic matter content and denitrification rates tended to be higher on older two-stage floodplains (Figure 11A), which suggests that N-removal capacity may depend on accumulation of organic matter over time. Organic matter may initially be low following floodplain construction, and then can accumulate through sediment deposition and via dynamics of floodplain vegetation. We found that both OM and denitrification rates at younger sites (CRE, KLA, CRO) tended to be higher in vegetated than in unvegetated soils (Figure 8D) and that denitrification rates were lowest during the first year after floodplain construction at CRE (Figure 11A) and at SHA (Roley et al., 2012a, b). In fact, previous research at SHA demonstrated a lag time of approximately one year before soil denitrification rates increased in response to addition of NO$_3^-$ rich stream water (Roley et al., 2012b), and we found that soil denitrification rates were nine times higher during the study period (2009-2010) than during previous research in 2008. However, soil organic matter content did not change significantly between 2008 and 2010, which could suggest a change in carbon quality or delayed development of microbial communities. Nevertheless, for denitrification, our results suggest that the two-stage ditch is a self-sustaining practice and may actually improve over time.

Denitrification rates on floodplain soils also varied seasonally with variation in NO$_3^-$ concentrations (Figure 7) and were N-limited up to 1 mg NO$_3^-$-N/L (Figure 9A), which suggests that floodplains had the capacity to respond to high NO$_3^-$ pulses during inundation but that denitrification efficiency was regularly saturated at high NO$_3^-$ loads. Priming could affect denitrification rates, which increased over the duration of a short inundation event at SHA, only becoming N-limited after four days (Roley et al., 2012b). This suggests that inundation patterns, including event frequency and duration, strongly affect NO$_3^-$ removal potential in floodplain soils (Baldwin and Mitchell, 2000).

**Estimating Reach-Scale N-Removal**

The two-stage implementation had the potential to increase N-removal at both base flow and during inundation by enhancing denitrification rates (Figure 7) and we found that aerial denitrification rates for mature floodplains (Figure 11A) were within the range of rates published for other restored wetlands (e.g., Mitsch et al., 2001, 2005). In scaling up our denitrification results based on areal denitrification rates and wetted widths for each site, we estimated that N-removal from the main channel of the stream.
ranged from 0.1 to 2.2 kg N/km/day (Figure 11C). Comparatively, during floodplain inundation, high denitrification rates combined with the three to five fold increase in bioreactive surface area could increase reach scale N-removal by 3-24 fold, with N-removal ranging from 1.8 to 10 kg N/km/day (Figure 11C).

Our monitoring of potential N removal by the two-stage ditch has great utility for managers and producers who often ask what the potential is for the two-stage to improve water quality. To that end, and to place our results in a nutrient management context, we use reach-scale denitrification estimates and a back-of-the-envelope calculation to estimate potential N-removal on an annual scale, for each km of two-stage implemented. Briefly, we estimated seasonal N-removal by multiplying reach scale N-removal (kg N/km/day) by the number of floodplain inundated days in each season, summing values for all seasons. We then compared these values vs. a rough estimate of annual NO$_3^-$ export which we based on seasonal means for NO$_3^-$ flux across our sites. Extrapolated N-removal via instream denitrification ranged from 22 to 630 kg NO$_3^-$-N/km/yr, which equates to 1-9% of annual N-export (Figure 12A). These estimates also suggest that, while two-stage floodplains likely had little effect on loads the first year after construction (CRE and SHA 2008, Figure 12C), mature floodplains could have reduced N-export by an additional 2-13% depending on inundation frequency (Figure 12B). Interestingly, Roley et al. (2012a) reported a similar percentage removal using much more detailed estimates, a different modeling approach, and a single stream for analysis. We note that these relatively small percentages are still significant in magnitude, given very high nitrate loads in these watersheds with intensive row-crop agriculture and extensive tile drainage. We also note that using our laboratory assays, we estimated high potential for N-removal via denitrification, but we did not quantify removal directly during inundation due to logistical constraints. Nevertheless, we argue that longer reaches could only improve the potential for detectable reductions in surface water NO$_3^-$ even with the very high loads seen in typical agricultural streams.

**CONCLUSIONS**

We evaluated the N removal capacity and water quality benefits of the two-stage ditch, in multiple agricultural streams where floodplains were constructed one to seven years prior to our study. At the reach-scale, the two-stage ditch significantly reduced surface water SRP concentrations and turbidity, and increased N-removal capacity via denitrification, but impacts on stream substrate composition and reductions in NO$_3^-$ loads were limited and/or stream specific. The two-stage ditch had the greatest impact on water quality when floodplains were constructed at a lower bench height, and function generally improved as floodplains matured, but effects were constrained by discharge and nutrient loads, which varied across watersheds and over time.

**ACKNOWLEDGMENTS**

Many thanks to A. Ward, J. Witter, K. Wamsley, and C. Watts for help with site selection and logistics. We thank private land
LITERATURE CITED


