Ecological Indicators 169 (2024) 112969

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

We know less about phosphorus retention in constructed wetlands than we think we do: A quantitative literature synthesis☆

Kenneth J. Anderson[a,*], Bishwodeep Adhikari[a], Olivia F. Schloegel[a], Raissa Marques Mendonca[a], Michael P. Back[a], Nicholas Brocato[a], Jacob A. Cianci-Gaskill[b], Steven E. McMurray[b], Christie Bahlai[a], David M. Costello[a], Lauren Kinsman-Costello[a]

[a] Department of Biological Sciences, Kent State University, Kent, OH 44242, USA
[b] Old Woman Creek National Estuarine Research Reserve, Huron, OH 44839, USA

ARTICLE INFO

Keywords:
Constructed wetland
Monitoring
Long-term monitoring
Phosphorus
Retention

ABSTRACT

Wetlands are increasingly being constructed to retain phosphorus, but indicators of performance are often unique to individual wetlands. In this quantitative literature synthesis, we highlight two major knowledge gaps and a source of bias in our current understanding of phosphorus retention by constructed and restored wetlands. We performed a literature review to assess differences in wetland characteristics, phosphorus retention, and sampling frequency and duration to better understand how constructed and restored wetlands retain phosphorus. We then examined a series of different sampling frequencies to determine how sampling approaches affect observed trends using a case study at the Old Woman Creek National Estuarine Research Reserve where high-resolution data on phosphorus concentrations was available. We found that, while differences among broad wetland structural groupings exist, measured rates of phosphorus retention are highly variable. Additionally, we observed that the wetlands sampled in the literature are most often newly constructed; 70% of the wetlands were monitored for three or fewer years. Among twelve wetlands that were monitored for over ten years, resampling techniques demonstrated that three-year time scales are unable to predict long-term trends for nutrient retention. Similarly, we found that 70% of wetlands in the literature were sampled weekly or less frequently and did not account for major flow events when estimating nutrient retention. Our case study analysis indicated that excluding storm events lead to a significant underestimate of phosphorus retention. We further demonstrate that when sampling is infrequent and misses storm events, the retention of water and phosphorus by wetlands causes higher underestimates at the inflow than the outflow where phosphorus leaves wetlands more slowly than the inflow. The greater underestimation of loads at outflows compared to inflows can lead to a wetland being mislabeled as a source when it is a sink for phosphorus if storm events are not accounted for. Together this synthesis highlights that our understanding of how constructed wetlands store phosphorus is limited by three major factors: a lack of long-term monitoring, the drivers of high variability in phosphorus retention between different wetland types, and potential bias from the difficulties of capturing storm events. Our analysis indicates that most published works have been conducted at shorter time frames, or lower sampling frequencies than are required for accurate estimates of phosphorus retention. This finding is of significance because substantial investments are made in restoring and constructing wetlands based on inadequately supported assumptions that wetlands are highly effective at phosphorus retention. Ultimately, these results can inform improvements to post-restoration assessment of wetland nutrient function, including decisions on where and when to monitor water quality.

☆ Open Research Statement: We have included data for the quantitative literature synthesis and resampling code for the case study as supplements for the manuscript. We have included all relevant citations, search terms and methods used to compile the data and include our derived datasets (literature search spreadsheets) and analysis code to ensure a reader would be able to assemble and identical dataset and repeat our approach. Data from the resampling case study at Old Woman Creek are available in the Environmental Data Initiative's Data Portal from McMurray et al (2024) at https://doi.org/10.6073/pasta/3f251395d82eafb33a93a9f3ebfb858c.

* Corresponding author.
E-mail address: kande120@kent.edu (K.J. Anderson).

https://doi.org/10.1016/j.ecolind.2024.112969
Received 3 September 2024; Received in revised form 8 November 2024; Accepted 4 December 2024
Available online 9 December 2024

1. Introduction

Excess phosphorus leads to eutrophication of aquatic habitats and fuels harmful algal blooms (Anderson et al., 2002; Heisler et al., 2008; Smith, 2003). A key challenge with managing phosphorus pollution is that it often does not originate from easily identified and mitigated "point" sources, but instead comes from "non-point" sources across wide swaths of landscape (Carpenter et al., 1998). To address this issue, scientists and land managers have proposed the creation and restoration of wetlands for their ability to store nutrients and specifically respond to the non-point source runoff of phosphorus (Vymazal, 2007; Weisner et al., 2016). Wetlands have three major pathways through which they may store phosphorus: sedimentation of suspended matter with phosphorus, sorption of phosphate to soil minerals, and uptake by vegetation (Reddy et al., 1999). Wetlands have the potential to be a cost-effective means of treating polluted waters, especially in the light of co-benefits such as the provision of habitat, carbon sequestration, and biodiversity (Aziz and Van Cappellen, 2021; Tomscha et al., 2021). However, an understanding of how wetlands contribute to landscape-scale nutrient retention requires a better understanding of the specific drivers of wetland phosphorus retention (Weisner et al., 2016). While a wealth of experimental evidence illustrates these storage mechanisms, demonstrating and predicting rates of uptake at whole-ecosystem scales in the field is not straightforward (Ury et al., 2023). Our current knowledge of how wetland features and their landscape context drive phosphorus cycling is not adequate to optimize wetland design and management for phosphorus retention.

Multiple efforts to synthesize data on wetland phosphorus retention have demonstrated that wetlands are generally phosphorus sinks and that higher retention of phosphorus is associated with higher phosphorus loading and higher flow rates (Golden et al., 2019; Land et al., 2016; Shen et al., 2023; Ury et al., 2023). A variety of wetland classification schemes are used (Gerbeaux et al., 2018), some with broad groupings based on wetland connectivity (Leibowitz et al., 2023), some that are highly granular considering multiple functional and structural aspects of wetlands (Cowardin, 1979), and some that group wetlands into broad categories, separating floodplain (low-lying areas next to rivers that are inundated during high flows), flow-through (defined as wetlands with a clear flow-path for water through the wetland), and isolated wetlands (wetlands that receive water primarily from overland flow), or constructed (man-made wetlands), restored (wetlands that have been modified in an attempt to restore natural function), and natural wetlands (wetlands that have not been modified; Tiner, 2015). Existing categorization frameworks, however, are less indicative of biogeochemical function than habitat type and offer limited insights into the diversity of structures found in novel restored and constructed wetlands.

Despite limitations of current formal classification schemes, certain prominent structural features, mostly relating to a wetland's hydrologic connection to its landscape, can be used to generalize conclusions about nutrient removal function across multiple systems described in the literature. For example, syntheses of data from river floodplains, demonstrate their value for nutrient retention, especially depressional wetlands within floodplains (Gordon et al., 2020; Walton et al., 2020). Nutrient retention by small and isolated wetlands have also been shown to support landscape-scale phosphorus transport and processing (Cheng and Basu, 2017; Marton et al., 2015). Research using mesocosms (small scale experimental wetlands) has provided unique insights into the mechanics of phosphorus retention, but mesocosms have been shown to have higher rates of retention than non-mesocosm wetlands (Ury et al., 2023). These syntheses have reaffirmed the potential value of wetlands for phosphorus mitigation, but also highlight that reliance on mesocosm studies overestimates wetland phosphorus retention capacities (Ury et al., 2023).

Attempts to use wetlands to mitigate excess phosphorus loads have had mixed success (Li et al., 2011; Steinman et al., 2018; Widney et al., 2018). At times, wetlands can be a source, rather than a sink, of phosphorus to downstream systems, often due to legacy stores of phosphorus (Ardón et al., 2010; Kinsman-Costello et al., 2014; Sharpley et al., 2013). Additionally, long-term data on constructed wetlands indicates that phosphorus retention is varied, with some wetlands declining rapidly in their ability to retain phosphorus and others maintaining their retention ability for decades (Kadlec, 2009; Steinman and Ogdahl, 2011). To improve use of wetlands for nutrient removal and retention, more information is needed on how their construction, management, and location within the landscape can maximize their role in nutrient mitigation (Djodjic et al., 2020; Schmadel et al., 2019; Stammel et al., 2018).

During construction, the placement of a wetland on the landscape determines the nature and magnitude of phosphorus inputs (Li et al., 2023). Flow to a wetland is a major factor in determining how much phosphorus a wetland can retain (Powers et al., 2013). High volumes of water entering a wetland can lead to high phosphorus retention, even when water retention is the primary goal of construction (Powers et al., 2013; Robotham et al., 2023). Despite the importance of wetland location, wetland placement is often determined by land availability as opposed to optimal nutrient load source treatment, so understanding if and how different wetland characteristics that affect their ability to retain phosphorus can be manipulated regardless of their position on the landscape is critical for the development of successful wetland projects (Locke and Grove, 2016; Turner et al., 2016). Here we define wetland characteristics as the structural and hydrologic attributes of a wetland, including vegetation status, hydrologic regime, water source, former land use, and wetland size.

Wetland physical structure influences hydrologic regimes, leading to the drainage, holding, or delivery of water to or from the system (Braskerud et al., 2000; Mitsch, 1992). After the physical characteristics, one of the most studied and significant aspects of wetland construction is substrate composition, which can be altered by the addition, removal, or movement of sediment and can modify the ability of sediment to bind phosphorus (Dell'Osbel et al., 2020). This is especially important where there is a significant amount of groundwater flow through the wetland; if there is sorption capacity in the sediment, phosphorus typically is stored there as water leaves through sub-surface upwelling or downwelling flow paths (Ackerman et al., 2015). Many wetlands are constructed with settling basins to encourage sedimentation and phosphorus storage in the accumulated sediment. However, as sediments build up wetlands may need to be dredged to maintain their ability to store water and phosphorus (Lindstrom and White, 2011; Yan et al., 1998).

Methods used to monitor nutrient loads have been widely studied in rivers and catchments, where data are abundant (Aulenbach et al., 2016; Birgand et al., 2010; Johnes, 2007; Kamrath et al., 2023; Kelly et al., 2018), but there has been limited study of the impact of sampling regime on estimates of nutrient loading in wetlands. Routine monitoring of a variety of rivers has allowed for the development of techniques to estimate phosphorus loading based on flow-concentration relationships; extensive evaluation of sampling regimes from high-resolution, distributed monitoring across many stream, river, and other catchment systems affirms that load estimates are most accurate when flow and concentration are measured over the full range of hydrologic conditions (Aulenbach, 2013; Aulenbach and Hooper, 2006). These methods are certainly relevant to wetlands as well but are limited in their use by the amount of data available for many wetlands. To better estimate phosphorus loading in wetlands we need to understand what techniques are currently being used, and potential biases in those techniques to develop wetland specific tools to deal with those biases.

Three critical pieces are missing from previous syntheses of net ecosystem phosphorus removal rates by wetlands: differences in retention caused by features of wetland construction, long-term (i.e., > 10 years) monitoring of phosphorus retention, and how differences in monitoring methods affect estimates of phosphorus retention. To address these gaps, we present a quantitative literature synthesis with the goal of addressing three main questions: 1. What attributes are most measured in wetlands, and does data collected to measure phosphorus retention performance vary by wetland structure? 2. How frequently and for how long have wetlands in the literature been monitored and how does that affect estimates of phosphorus retention? 3. Are temporal trends of wetland phosphorus retention within the confines of typical study lengths able to predict the future trajectory of wetland phosphorus retention? To evaluate wetland attribute information and monitoring methods, we conducted a quantitative literature review. To assess the implications of sampling regime on P load estimates, we resampled high-frequency data from a single wetland case study. Finally, we analyzed a subset of wetlands from our quantitative review with annual data on phosphorus retention for long-term trends. We aim to identify gaps in our understanding of structural wetland features, monitoring methods, and long-term trajectories in restored and constructed wetlands.

2. Methods

2.1. Literature review

We conducted a systematic search of peer-reviewed publications focused on phosphorus retention by constructed or restored freshwater wetlands that contained at least one year of data estimating annual total phosphorus and/or phosphate (PO_4^{3-}) retention. We excluded all mesocosm studies, which were defined following Ury et al. (2023) as studies conducted in tubs, tanks, or other semi-isolated structures. While there are differences between constructed, restored, and natural wetlands, we aim to capture the differences between different wetlands by characterizing the different attributes that differentiate wetlands as opposed to their construction history. We included a variety of wetland types, including flow-through, riparian, and managed wetlands but excluded studies that were focused on streams, vegetated buffer zones, or bioreactors. Streams and vegetated buffer zones both play an important role in phosphorus retention across the landscape but function differently than wetlands, which have periods with standing water (Uusi-Kämppä et al., 2000).

We performed an initial search using Web of Science and then supplemented the list with the first 300 results from Google Scholar; we then included all manuscripts used by Ury et al. (2023) that had not appeared in our Web of Science or Google Scholar searches. For both Web of Science and Google Scholar searches we followed the approach of Ury et al. (2023) for keywords and searched with the following terms: *(wetland * OR pond OR mire * OR marsh OR fen OR 'wet meadow' OR riparian OR 'floodplain' OR reed) AND (phosph *) AND (retention OR trap * OR uptake OR sedimentation OR remov * OR settling OR accretion OR precipitat * OR * sorption)*. Our combined search yielded 4,170 unique results ranging from 1997 to 2023. We performed an initial screening of titles for papers that were relevant to wetlands which narrowed our results to 975 studies, and then a second screening for papers with at least one year of data on phosphorus retention, measured as a load of phosphorus. The second screening narrowed our results down to 207 manuscripts. For 48 studies, we calculated loads based on the reported flow and concentrations at wetland inlets and outlets.

For each wetland, we extracted information about both the structure and the monitoring methods used if they were available, and we classified wetland vegetation, flow regime, water source, prior land use, phosphorus sampling frequency, gap-filling methods, age of wetland, and length of monitoring. Individual wetland characteristics were then used for grouping analyses and study monitoring characteristics were used to understand monitoring trends. A subset of wetlands with phosphorus retention data for ten or more years was used to analyze long-term trajectories of wetland phosphorus retention.

2.1.1. Wetland structure

There was a limited amount of available information about each wetland in the literature so we used the Cowardin classification system, an ecologically based framework developed to provide a consistent classification system across US wetlands (Cowardin, 1979; Cowardin and Golet, 1995) to classify each wetland system as riverine, lacustrine, or palustrine. Wetlands may meet or partially meet the criteria for classification into multiple system types, so we developed a set of decision rules for these borderline cases. We classified any wetlands within channel floodplains as riverine, whereas wetlands that were periodically flooded but not within a regular floodplain as palustrine. Similarly, we classified studies as lacustrine only if the retention of phosphorus was measured for the entirety of a lake that contained many coastal wetlands; studies focused on estimating retention at a single wetland on the shore of a lake were considered as palustrine. For wetland class (i.e., emergent, aquatic bed, etc.), modifiers (i.e., beavers, wastewater, etc.), water chemistry (i.e., acidic, basic, neutral), soil (i.e., organic, inorganic), and water regime (i.e., seasonally flooded, intermittently exposed, etc.) we followed Cowardin classification standards (Cowardin, 1979). We additionally recorded the source of water, the previous land use, and the watershed area when reported, as well as the surface area and water depth of the wetland. We classified the water regime as artificial when actively managed, most commonly when water was pumped in or managed using a water-level control structure.

2.1.2. Wetland monitoring

We extracted information on the methods and extent of monitoring for each wetland. First, we classified whether estimates of phosphorus retention were based on sediment or water measurements. Then, the frequency of both nutrient and flow sampling was recorded. Finally, we noted whether surface water was sampled during storm events and how studies accounted for gaps between surface water concentration data points. We identified four common methods of gap-filling and grouped any techniques used only in a single manuscript as "other." The four main methods were 1. Assuming constant, where concentrations were assumed to be constant until a new measurement is taken, 2. Flow-weighted means, where the mean annual concentration was calculated using weekly or monthly concentrations adjusted based on discharge and then the annual mean concentration is used as the concentration for each day of discharge, 3. Linear interpolation, where a linear model was created to fill in gaps between data points, and 4. Event sampling, where concentrations are assumed to be constant except during events, when samples are taken. Studies where a load was not calculated (where we calculated loads based on concentration and flow data) were not included in identifying a gap-filling method.

2.1.3. Long-term trajectory analysis

To estimate the amount of data needed to identify long-term trends, we used the broken window algorithm developed by Bahlai et al. (2021), which splits time series into all possible sequential subsets of at least three years and constructs linear regressions of each subset for a variable of interest: in our case, phosphorus retention. We then calculated the stability time as the window length at which all regression slopes were within the standard deviation of the slope of the full dataset (Fig. S1, Bahlai et al., 2021). We applied the broken window algorithm to the twelve wetlands from six manuscripts that contained an annual measurement of phosphorus retention for ten or more years.

2.1.4. Cluster analysis

We performed a grouping analysis to avoid any bias in our grouping based on any specific method. To identify the major wetland groupings that emerged from the recorded attributes, we calculated Gower's distances based on seven wetland features: Cowardin wetland system, Cowardin wetland class, wetland water regime, water source, previous land use, whether the water regime was artificial or not, and the surface area of the wetland. After calculating Gower's distances, we performed k-means clustering with two to ten clusters and selected the number of clusters that maximized the silhouette width. The silhouette width is a metric of how similar objects are to their clusters as compared to other clusters. The calculations of Gower's distances and the cluster analysis were performed using version 2.1.6 of the *cluster* package in R (Maechler et al. 2023). We compared differences between groups with analysis of variance and Tukey HSD post-hoc tests. To visualize the cluster data, we used the t-distributed stochastic neighborhood embedding (tSNE) method from version 0.17 of the *Rtsne* package (Krijthe, 2014) to reduce the dimensionality of our wetland data and plot points in a two-dimensional space that preserves the differences between our individual wetlands. tSNE plots arrange data so that points with large differences in high dimensional space are far apart, and points that are similar in high dimensional space are close together in the two-dimensional plot.

2.2. Wetland case study: Old woman creek

We resampled high-frequency, long-term data from the Old Woman Creek (Ohio, USA) wetland to illustrate how changes to sampling frequency, gap-filling, and storm sampling affect estimates of phosphorus retention. The Old Woman Creek National Estuarine Research Reserve is a coastal wetland estuary that drains into southwestern Lake Erie. Old Woman Creek is a fourth order stream with a 69 km^2 watershed dominated by row crop agriculture. The wetland is a natural estuary, but it receives high nutrient loading from its agriculture dominated watershed. The wetland is a mix of emergent vegetation and aquatic bed, and a majority of the site is permanently flooded. The final 2 km of stream flows through a 0.6 km^2 wetland complex before flowing into Lake Erie, although outflow is controlled by a barrier sand beach. Generally, the barrier beach is closed during the summer and open throughout winter, although its status varies from day to day and year to year.

Previous efforts have estimated phosphorus budgets for Old Woman Creek (Mitsch and Reeder, 1992; Krieger, 2003), and quantifying the wetland's capacity for processing nutrients remains an important line of inquiry. Previous nutrient budgets have relied on inlet water flow measured by the United States Geological Survey (USGS; gage 04199155). However, the installation of an Acoustic Doppler Current Profiler at the stream's outlet in 2020 has improved the accuracy of current estimates of phosphorus loading into Lake Erie. Water flow, which is measured at fifteen-minute intervals at the inlet of Old Woman Creek, is available throughout the study period from the USGS gage, while water flow at the outlet is intermittently available beginning in May of 2020 and full-time beginning in 2023.

Starting in 2015, whole water samples were collected with ISCO 5800 refrigerated autosamplers at both the inlet and outlet of the wetland for total phosphorus analysis. From 2015 to 2019, daily water samples at both the inlet and outlet sites were taken at 04:00, 12:00, and 20:00. During baseflow, a single sample was analyzed for each day, and during storm events three samples were analyzed at both inlet and outlet. An analysis of 2015–2018 nutrient concentrations showed reduced temporal variability at the outlet compared to the inlet. During baseflow, concentrations remained stable over the course of a week, and, during stormflow, concentrations varied from day to day, but not throughout the course of a single day (Johnson, 2018). Therefore, beginning in 2019, sampling frequency at the outlet was reduced to once per week during baseflow, and once per day during stormflow. At the inlet, the original sampling frequency of once per day during baseflow and three times per day (04:00, 12:00, 20:00) during stormflow has continued since 2015.

All water samples were measured for total phosphorus by the Old Woman Creek Reserve's analytical lab. Analysis started with a persulfate digestion on a hot plate before samples were analyzed with an ascorbic acid reduction following Standard Methods 4500-P) (Baird et al., 2017).

2.2.1. Resampling methods

For our resampling analysis, we focused on sampling frequency scenarios for phosphorus concentrations that reflect the common sampling schemes and gap-filling observed in our literature review. We used all available daily measurements of flow across all analyses because continuous flow data are more likely to be available via USGS gages or other sensors. The only variable that changed across our "simulations" was phosphorus concentrations. We considered the estimate based on all available daily concentration data as the best possible estimate of "true" phosphorus loading. We created functions in R to resample phosphorus concentrations from the full dataset at weekly and monthly frequencies. For each variation, we wrote a function to select a random day of the week or month to use as the monthly value using a combination of base R, *dplyr*, *tidyr*, and *tidyverse* packages (R Core Team, 2024; Wickham et al., 2023, 2019; Wickham and Wickham, 2017). We tested gap-filling using four different methods: constant, flow-weighted means, event sampling, and linear interpolation. Constant refers to an assumption that concentrations remain constant between sampling points. Event refers to an assumption that concentrations remain constant between sampling points, except during storm events. The flow-weighted approach calculates flow-weighted mean concentrations to calculate the mean concentration over time. The linear interpolation approach computes a linear model between points and fills in gaps using the linear model. For the constant method, we applied the monthly concentration to all days of that month. For flow-weighted means, we calculated the flow-weighted average concentration for the year from the weekly or monthly sampling, and then applied that single value to every day in the dataset. We calculated the flow-weighted average concentration as the sum of the weekly or monthly concentrations multiplied by the percent of annual flow that each week or month accounted for. For linear interpolation, we gap-filled between monthly concentrations using the *approx* function in R.

We additionally resampled scenarios at monthly sampling frequency with the addition of storm samples. We defined storms as days with a discharge at the USGS stream gage of $>$ 25 cubic feet per second and large storms as days with a discharge at the inlet of $>$ 100 cubic feet per second. We selected 25 cubic feet per second because observed baseflow during the monitoring period was never greater than this threshold, making it easy to attribute elevated flow to storm events, and discharge has been observed to exceed 25 cubic feet per second in every season. An individual storm at 25 cubic feet per second was on average 1 % of the annual inflow to the wetland. We selected 100 cubic feet per second to limit analyses to only the largest storms and keep the number of yearly storms below 30. An individual storm at 100 cubic feet per second was on average 4 % of the annual inflow to the wetland. We resampled the data record by randomly selecting between one and ten storm events, for both all storms and only large storms. We then applied the average stormflow concentration to all unsampled storms or large storm events, respectively.

2.2.2. Uncertainty analysis

When comparing our resampling methods, we assumed that our daily data with increased sampling frequency during storm events is the best possible estimate of the "true" load as this sampling frequency is realistically closest to continuous measurements of phosphorus concentrations (Aulenbach et al., 2016; Birgand et al., 2010; Kamrath et al., 2023; Walling and Webb, 1981). For each resampling method, we replicated the selection process one thousand times and then compared them to our assumed "true" loading data.

We calculated the relative uncertainty of each year of sampling replicate as e (%) = (estimated load − true load) / true load) * 100. We then calculated the bias for each year of each sampling method as the median e, and the imprecision as: imprecision = e90 − e10. Where e90 is the 90th percentile of e, and e10 is the 10th percentile of e (Appling et al., 2015; Birgand et al., 2010; Kamrath et al., 2023).

3. Results

3.1. Wetland types

Our literature analysis included 415 constructed or restored wetlands (from 207 studies) with phosphorus budgets across all continents except Antarctica, with a bias for the United States and Northern Europe (Fig. 1). In our grouping analysis, silhouette width reached an initial peak at four clusters and then continued to increase after six clusters (Fig. 2). This pattern highlights the high diversity in different types of wetlands (based on Cowardin wetland system, Cowardin wetland class, wetland water regime, water source, previous land use, water regime, and surface area) as, after being grouped into four clusters, individual wetlands became increasingly similar to smaller distinct groups. The four groups identified were primarily split by inflowing water source and land use setting: wastewater treatment (characterized by wastewater as a water source), agricultural runoff (characterized by agricultural land use), stream/river water (characterized by streams or rivers as a water source), and floodplain wetlands (characterized as part of riverine systems) as the major groups (Table S1). The group we classified as floodplains were intermittently flooded, compared to wetlands receiving stream water which had standing water for longer periods. Wastewater treatment wetlands retained the most phosphorus (ANOVA, $F_{(3,309)} = 4.66$, $P < 0.001$; Table 1), specifically retaining more phosphorus than wetlands treating agricultural runoff (Tukey HSD, $P < 0.001$) or wetlands treating stream water (Tukey HSD, $P < 0.05$). Similarly, wastewater treatment wetlands received the most phosphorus loading (ANOVA, $F_{(3,297)} = 7.94$, $P < 0.001$; Table 1), with more loading than wetlands treating agricultural runoff (Tukey HSD, $P < 0.001$), wetlands treating stream water (Tukey HSD, $P < 0.001$), or floodplain wetlands (Tukey HSD, $P < 0.05$). For all wetland types, the variability in both phosphorus retention and phosphorus loading was extremely high, with phosphorus retention ranging from −140 to 1500 g/m^2/year for an average of 51 ± 164 g/m^2/year. Phosphorus retention in wastewater treatment wetlands was 100 ± 201 g/m^2/year (average ± standard deviation), in wetlands treating agricultural runoff was 15 ± 37 g/m^2/year, in wetlands treating stream water was 35 ± 137 g/m^2/year, and in floodplains was 69 ± 301 g/m^2/year. In all wetland categories the standard deviation was greater than double the average phosphorus retention.

3.2. Wetland sampling frequency

We found that 138 of the 207 studies constructed hydrologic budgets, and among those studies 59 did not report the frequency of hydrologic sampling. Of the remaining 78 studies, 64 had continuous measurements of flow, 10 actively pumped water in at a consistent rate, 1 substituted precipitation for flow, and 2 measured flow weekly. The majority of studies reported constructed hydrologic budgets with high frequency data, however, the frequency at which phosphorus concentrations were sampled varied far more widely between studies. Across the 207 included studies, 138 used water samples to measure phosphorus retention and 22 used measures of sediment accretion to measure phosphorus retention. Of the 138 that used water samples, the most common phosphorus sampling frequency was weekly (34 studies) followed by monthly (31 studies). Daily data was collected in 22 studies and was predominately collected using an autosampler. Unknown phosphorus concentrations between sampling points were estimated by assuming that concentrations were constant for 57 % of the studies (Fig. 3). The most common alternate methods of gap-filling were flow-weighted means and linear interpolation, which combined only accounted for 11 % of the studies that used water samples. Similarly, event samples were collected in only 9 % of the studies using water samples.

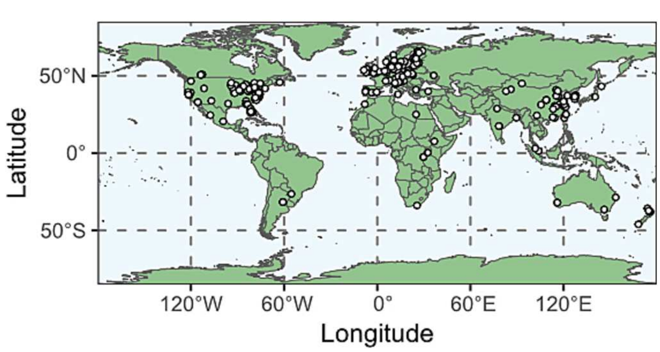

Fig. 1. Locations of the 415 wetlands included in the literature synthesis; all have reported phosphorus retention for at least one year.

3.3. Wetland sampling duration

We found a distinct paucity of long-term monitoring of wetlands for net phosphorus budgets. Of wetlands with phosphorus retention data reported in the literature, 70 % were sampled for three or fewer years (Fig. 4a). The majority (51 %) of data available were from wetlands two or fewer years old (Fig. 4b).

Of the 33 wetlands that monitored phosphorus retention for ten or more years, 12 had data available on annual averages of phosphorus retention. Of the remaining 21 wetlands, 13 measured phosphorus retention using sediment data, and eight reported an average across the period monitored. Of the 12 sites with long-term data, six were expansive Everglades stormwater treatment wetlands treating agricultural runoff, four were small Northern European wastewater treatment wetlands treating sewage, one was a Northern European peat extraction wetland, and one was a shallow lake in Michigan used to treat wastewater (Table 1). Out of the 12 sites, four had information about dredging, and, of those four, one was dredged and three were not. Notably, long-term data is available for a wider variety of wastewater treatment wetlands, while long-term data on wetlands treating agricultural drainage only include Everglades stormwater treatment areas. All 12 wetlands were still retaining phosphorus at the end of the sampling period with a range of 4 to 1135 g P/m^2/year and a range of 10 % to 99 % retention of incoming phosphorus in the last year of sampling (Table S1). Eight of the wetlands were monitored monthly, two bimonthly, one biweekly, and one seasonally, and none of them accounted for storm sampling (Table S1).

Three of the twelve sites had statistically significant increases in phosphorus retention over time, two had statistically significant decreases in retention over time, and seven were stable over the sampling period. The sites were sampled for an average (±SD) of 15 ± 5 years and had an average (±SD) stability time of 9 ± 3 years (Table 1). We looked specifically at the accuracy of sampling for long-term trends after five years of monitoring because we found that data on phosphorus retention of wetlands sharply drops off after five years (Fig. 4). The five-year windows created with the broken window algorithm were directionally incorrect (ex. Positive trends were instead negative over longer time scales) on average 45 ± 37 % of the time (Table 1).

It is important to note that the stability time is relative to the total number of years monitored. We found a strong relationship between the stability time and the number of years monitored (Fig. 5). The trajectory of this trend past twenty years is largely driven by a single point, but it suggests that, for long-term data, you can predict phosphorus retention trajectories for 1.6 ± 0.1 (slope ± standard error) years in the future for every year you monitor after at least five years of monitoring. When the furthest timepoint at thirty years is removed, slope changes to 1.5 ± 0.4 (slope ± standard error). This additionally highlights the need for data on wetlands monitored for longer than twenty years.

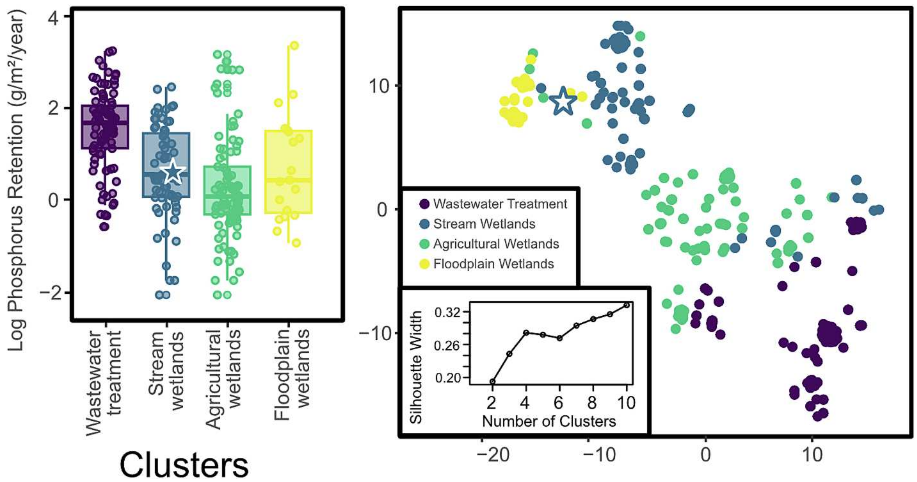

Fig. 2. Boxplots illustrating the log phosphorus retention and a t-distributed Stochastic Neighbor Embedding (tSNE) plot for 4 clusters (the number of groups where the silhouette width peaks before becoming increasingly granular). Dissimilarity between wetlands was calculated using Gower's distance, and tSNE plots were constructed to visualize similarities. tSNE plots visualize high dimensional data in two dimensions using machine learning to position points with large differences in high dimensional space far apart in the two-dimensional plot, and points that are similar in high dimensional space are close together in the two-dimensional plot. Boxplots represent the interquartile range of log phosphorus retention for all wetlands within a cluster, the solid lines are the median, error bars represent the 95% confidence interval, and jittered points represent the actual datapoints. Groups were delineated by k-means clustering to 4 clusters based on the peak in silhouette width, before clusters become increasingly granular. Clusters were named based on the attributes that were dominant in each cluster (see Table S1). The star in both plots represents Old Woman Creek which we use as a case study.

Table 1
Broken window statistics for sites with > 10 years of data on phosphorus retention. Stability time indicates the length of sampling it takes to have all estimates of the long-term slope within the standard deviation of the full-time series.

Years Monitored	Stability Time (Years)	Proportion Wrong After 5 Years	Long-Term Slope	Reference
12	8	0.83	−0.19	(Pietro and Ivanoff, 2015)
12	7	0.83	−0.16	(Pietro and Ivanoff, 2015)
12	7	0.25	−0.01	(Pietro and Ivanoff, 2015)
12	8	0.12	−0.08	(Vymazal 2011)
13	7	0.11	−0.02	(Ronkanen and Kløve, 2009)
14	9	0.8	0.13	(Jóźwiakowski et al. 2018)
15	9	0.18	−0.02	(Ronkanen and Kløve, 2009)
16	9	0.1	0.03	(Pietro and Ivanoff, 2015)
16	8	0.16	−0.03	(Pietro and Ivanoff, 2015)
16	10	0.16	−0.02	(Pietro and Ivanoff, 2015)
18	10	1	0.1	(Krzeminska et al., 2023)
30	18	0.83	0.05	(Kadlec, 2009)

3.4. Case study: Old woman creek

Phosphorus sampling frequencies at weekly and monthly intervals both lead to a significant negative bias in annual retention estimates for wetlands (Fig. 6). Weekly sampling was on average 40 % negatively biased compared to daily values, while monthly sampling was nearly 60 % negatively biased. At Old Woman Creek the annual calculation of phosphorus loading using weekly sampling was on average 2,500 lbs (1,100 kg) less than values calculated with daily sampling and on average 3,000 lbs (1,350 kg) less than daily sampling when using monthly sampling. The inclusion of storm event sampling reduced negative bias but also increased imprecision compared to monthly and weekly sampling (Fig. S2). Preferential sampling of large storms was both better at improving bias than small storms and less imprecise than small storms (Fig. S2). There was no significant difference in bias based on the number of storms sampled (Fig. S2); however, sampling more storms led to less imprecise data (Fig. S2). Calculating phosphorus loading with flow-weighted means or linear interpolation did not improve bias or imprecision of estimates (Fig. S2). Sampling storm events significantly improved bias estimates and prioritizing large storm events improved bias estimates even further (Fig. 6). Total phosphorus concentrations at the inlet and outlet of Old Woman Creek averaged 0.11 (median 0.08) and 0.13 (median 0.10) mg/L between 2015 and 2023, respectively; however, the variation in phosphorus concentration at the inlet was double the variation at the outlet throughout the data record (Table 2). Between 2015 and 2023, we found a significant positive relationship between discharge and P concentration at both the inlet ($R2 = 0.26$, $t = 33.3$, $P < 0.001$) and the outlet ($R2 = 0.08$, $t = 10.72$, $P < 0.001$) of the wetland.

We additionally found that there was a significant difference in bias between wetland inlets and outlets when both are sampled monthly (Tukey HSD, $P < 0.05$; Fig. 6) and weekly (Tukey HSD, $P < 0.05$; Fig. 4). Imprecision was lower at the outlet than the inlet, but not significantly when sampled either monthly (Tukey HSD, $P = 0.23$; Fig. 6) or weekly (Tukey HSD, $P = 0.93$; Fig. 6). The difference in bias between inlets and outlets caused negative average estimates of wetland phosphorus retention when sampling frequency was decreased from daily to monthly for multiple years studied (Fig. S3). This shifted estimates of the wetland from a sink of phosphorus to a source, based only on changing phosphorus sampling frequency from daily to monthly sampling (Fig. 7).

4. Discussion

4.1. Trends in wetland phosphorus retention

Measurements of phosphorus retention by constructed and restored wetlands in the literature are focused on a highly variable but non-representative subset of wetland types, are likely biased, and do not indicate long-term trajectories. Our synthesis underscores that wetlands are highly diverse and that broad categories based on traditional

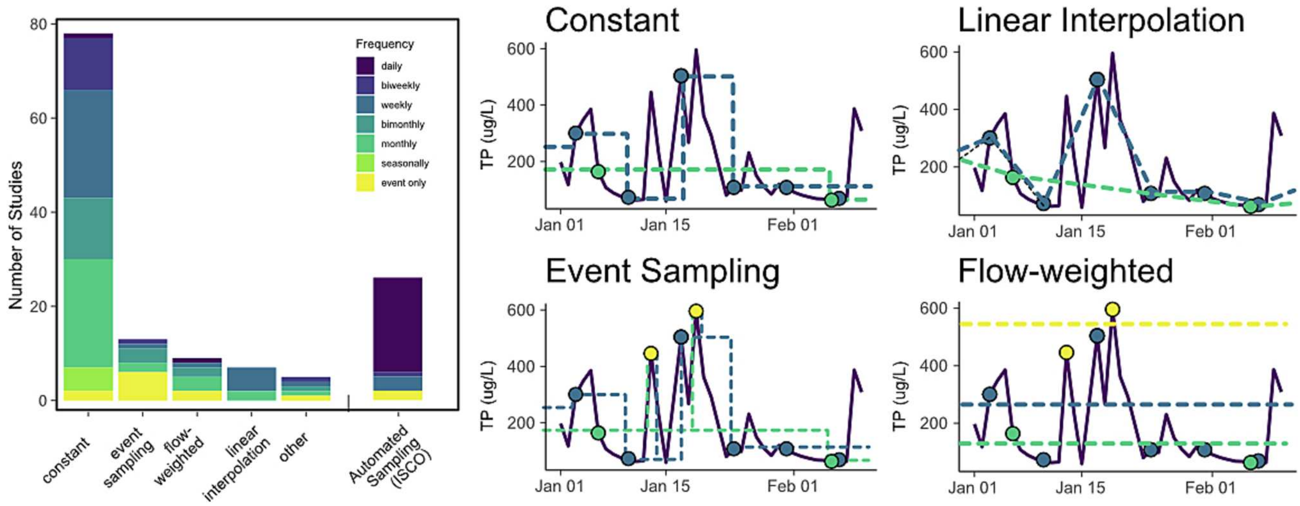

Fig. 3. Gap-filling methods for a wetland sampled at different sampling frequencies compared to sites with "continuous" ISCO collected flow-weighted composite phosphorus concentrations and examples of what those gap-filling methods look like in a real dataset at monthly, weekly, and event sampling time frames. Colors indicate frequency on both sides of the figure. Dashed lines represent either weekly (in blue) or monthly (in green) sampling, while the solid purple line is the daily data that weekly and monthly variations are derived from. Each gap filling approach is described in the methods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

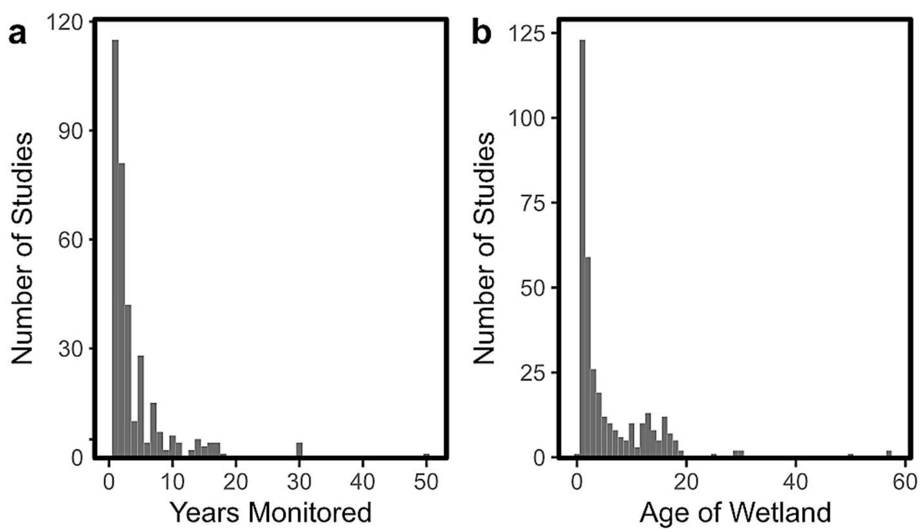

Fig. 4. Histograms highlighting a) the number of years monitored and b) the age of the wetlands studied across the 201 studies used.

structural features are insufficient to explain variability in P function. Research to date has established the basic mechanisms by which plant uptake, sediment storage, and soil phosphate sorption can retain phosphorus within wetlands (Reddy et al., 2022). However, mechanistic understanding of soil- and patch-scale processes are challenging to scale up to whole-system predictions. Wetlands are spatially heterogeneous, have multiple complex features, often have managed hydrology, and particularly in the case of novel constructed and restored ecosystems are designed for a variety of purposes (Hambäck et al., 2023). Our synthesis reveals the dearth of whole-ecosystem wetland P budgets that integrate net effect of heterogenous processes. Of the few whole-system P budgets in restored and constructed wetlands, the diversity in the kinds of wetland systems represented is inadequate and few have been monitored for long enough and at an appropriate frequency. The acquisition of accurate data that are representative of a wide range of wetland types is increasingly important to quantify the efficacy of wetlands created to retain nutrients (Djodjic et al., 2020; Ury et al., 2023) and informs costly investments in wetlands to mitigate nutrient pollution. Investment in wetland construction is increasing, presenting an opportunity to better understand how artificial ecosystems function and contribute to the retention of phosphorus (Clifford and Heffernan, 2018).

While we found similar trends to past analysis, that wetlands with higher phosphorus loading also have higher retention (specifically wastewater treatment wetlands), phosphorus retention was highly variable within wastewater treatment wetlands (Ury et al., 2023). Water source plays an important role in phosphorus loading and retention, however the unexplained variability between wetlands with the same water sources highlights the impact of other wetland features such as flow regime and vegetation, which contribute to the increasingly smaller clusters in our analysis. These smaller clusters combined with the high variability within broad clusters indicate that current literature estimates are insufficient for assessing the drivers of phosphorus retention in constructed wetlands. There are many factors that we know to be consequential for wetland P retention based on mechanistic studies that were not adequately reported across our synthesized studies to be included in our analysis-including vegetation, pH, and sediment composition.

Wetland vegetation can take up and store phosphorus and transform

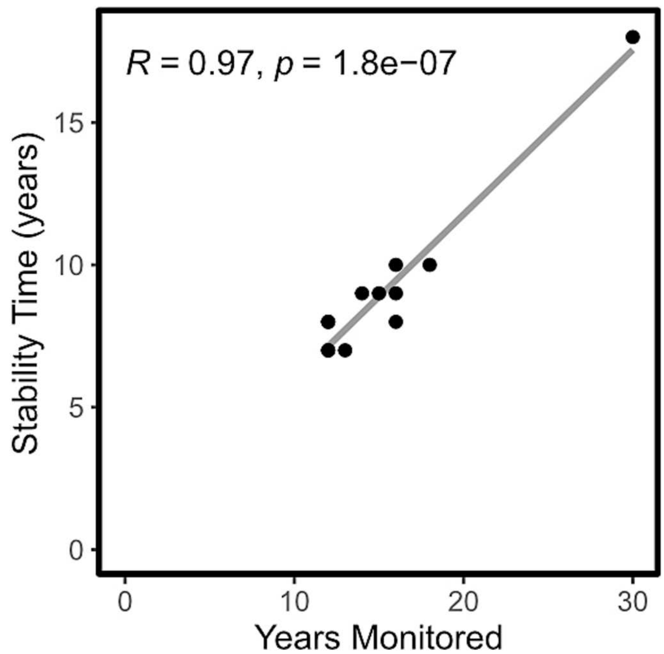

Fig. 5. Linear relationship between stability time and years monitored for 10 wetlands with greater than 10 years of data on phosphorus retention. For each wetland, the time series was split into all possible sequential subsets of at least 3 years and the linear regression of phosphorus retention over time was calculated. The stability time was calculated as the subset length at which all regression slopes were within the standard deviation of the full dataset (i.e., when the time series was long enough to predict the result of the full dataset).

inorganic phosphorus to more stable organic forms (White et al., 2004; Kiedrzyńska et al., 2008). While we show that there are distinct groups of wetlands categorized by vegetation, there is a huge variety of vegetation species compositions present in wetlands that we were unable to account for in this synthesis, because details on vegetation composition were either not reported or inconsistently reported throughout the literature. This data gap highlights the need for more study on how vegetation composition affects phosphorus retention in constructed wetlands. Similarly, variability in the flow regime of constructed wetlands has been shown to have large effects on their ability to retain phosphorus (White et al., 2004; Zhang et al., 2011). As we show with our grouping analysis, our ability to draw conclusions about how variable flow affects phosphorus retention is complicated because groups defined by flow variability also have major confounding factors (such as water source, and vegetation patterns). Furthermore, we show that the existing literature data may be biased as a consequence of changing flow patterns (storm events), which may have a bigger effect on certain flow regimes compared to others.

4.2. Phosphorus sampling in published literature

A variety of literature using stream data highlights the potential biases in estimating nutrient loading based on infrequent sampling for solute concentrations and provides tools for improved loading estimates (Appling et al., 2015; Aulenbach et al., 2016; Johnes, 2007). We found in our literature synthesis that 72 % of studies that have calculated nutrient loads into and out of wetlands did not account for potential biases related to storm events due to infrequent phosphorus sampling, which we show can lead to inaccurate estimates of phosphorus retention; the studies that accounted for bias related to storm events did so through event sampling (9 %) and daily sampling of phosphorus concentrations (19 %). The issue of inaccurate estimates of phosphorus loading is further exacerbated by the fact that wetland studies over shorter time scales are likely not able to predict long-term changes to wetlands. Yet, a majority of documented wetlands among our literature set were sampled for less than three years, which, at low frequencies of sampling, makes many of the methods used to calculate stream loading, such as regression models, less effective (Aulenbach, 2013; Aulenbach and Hooper, 2006). Rather than use ineffective models, we elected to

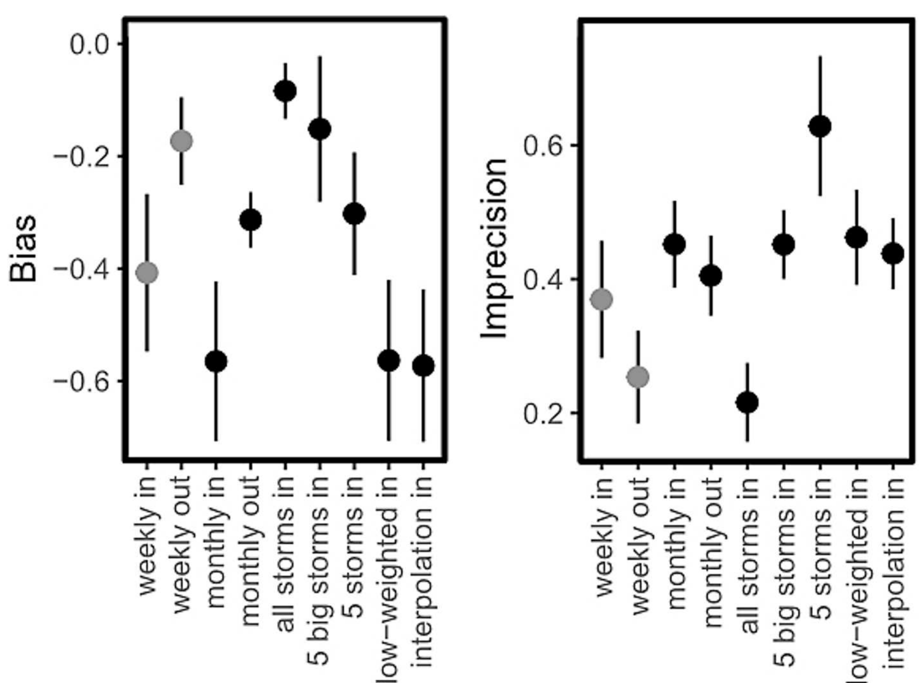

Fig. 6. Bias and imprecision of estimates of phosphorus loading at Old Woman Creek across a range of resampling scenarios. Bias is calculated as the median error of the 10,000 replicates. Imprecision is calculated as the 90th percentile of the error minus the 10th percentile of error. The standard deviation for both metrics is calculated as the variability between the nine different years (2015–2023) in the dataset. Sampling scenarios for weekly (grey points) and monthly (black points) are described in the methods.

Table 2
First, second, and third quartiles of total phosphorus (TP) concentrations (mg/L) at the inlet and outlet of Old Woman Creek wetland from 2015 to 2023. Between 2015 and 2019 at both the inlet and outlet, one and three samples per day were collected during baseflow and stormflow, respectively. Sampling frequency at the site changed beginning in 2019, when sampling at the outlet was reduced to once per week during baseflow, and once per day during stormflow.

Year	First Quartile (25 %) inlet TP (mg/L)	Second Quartile (50 %) inlet TP (mg/L)	Third Quartile (75 %) inlet TP (mg/L)	First Quartile (25 %) outlet TP (mg/L)	Second Quartile (50 %) outlet TP (mg/L)	Third Quartile (75 %) outlet TP (mg/L)
2015	0.04	0.06	0.08	0.07	0.10	0.15
2016	0.03	0.06	0.09	0.07	0.10	0.15
2017	0.04	0.06	0.10	0.09	0.12	0.15
2018	0.04	0.07	0.11	0.09	0.13	0.18
2019	0.05	0.08	0.14	0.08	0.11	0.15
2020	0.05	0.10	0.14	0.10	0.13	0.22
2021	0.09	0.12	0.18	0.10	0.19	0.27
2022	0.08	0.11	0.15	0.13	0.16	0.20
2023	0.07	0.10	0.15	0.13	0.17	0.26
2015–2023	0.05	0.08	0.13	0.08	0.12	0.17

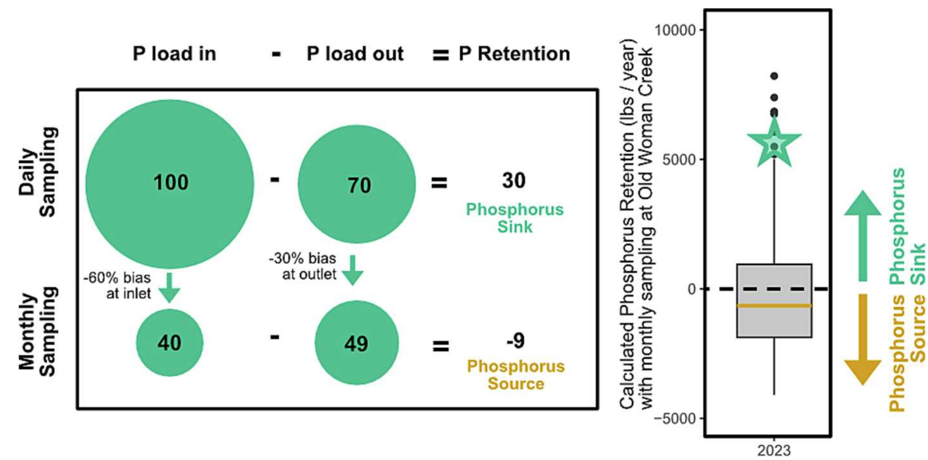

Fig. 7. Conceptual diagram illustrating that a mismatch in bias at the inlet and outlet of wetlands can lead to an inaccurate calculation of the phosphorus retention status of wetlands, accompanied by resampled data from Old Woman Creek illustrating this happening in a real system. Bias estimates in the conceptual diagram approximate the average biases we measured when converting from daily to monthly sampling. Resampling estimates from Old Woman Creek were resampled at a monthly frequency from daily data. The boxplot represents the interquartile range of the 10,000 replicate resamples of the original daily data, the solid line is the median, error bars represent the 95% confidence interval, and black points represent outliers. The green star indicates the true phosphorus retention value for the year calculated using daily data. This pattern was evident in all years measured (Fig. S3). Discharge data was not resampled and was constant across all resampling replicates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

use resampling to highlight where specific sources of bias come from, so that researchers can identify and avoid biased data collection.

Our resampling case study indicates that phosphorus concentration measurements during large storm events significantly improve the accuracy of annual phosphorus retention estimates from infrequent sampling, supporting past studies highlighting this same approach (Aulenbach et al., 2016; Richards and Holloway, 1987; Zhu et al., 2012). Furthermore, in Old Woman Creek we noted a mismatch in bias between inlet and outlet loading, where the loading values at the outlet were significantly less variable. This led to higher bias (in this case higher underestimates) in calculated phosphorus retention at the inlet compared to the outlet. The higher underestimate of P load at the inlet than the outlet is likely caused by the retention of water, and thus phosphorus by wetlands. Wetlands slow water flow, increase hydrologic retention time, and often smooth out peaks of flow (Braskerud et al., 2000; Ghosh and Gopal, 2010; Mitsch, 1992). The retention of water also contributes to the retention of phosphorus, which likely explains the lower variability in phosphorus concentrations at the outlet, where nutrients are slowly released after being stored in the wetland (Persson et al., 1999). Storm events that bring large loads of phosphorus during a single day are unlikely to be captured at the inlet, but the release over multiple days at the outlet is more likely to be captured. This slow release of phosphorus is especially important for wetlands in which the calculated percent phosphorus retention is lower than the difference in bias between the load at the inlet and the outlet. Retention is calculated as the difference between inlet and outlet loads, which means that with a mismatch in bias, a wetland can appear to be a source of phosphorus when it is actually a sink, where storms are more likely to be captured at the outlet than the inlet (Fig. 7). This finding has fundamental implications for how sampling must be designed to ensure accurate estimates of wetland function.

Our case study represents only a single wetland with a large capacity to store water and highly storm-driven hydrology from an agricultural watershed. It does however highlight how wetlands sampled infrequently and with high phosphorus concentrations during storm events (that is effectively captured by the wetland) are likely to have similar biases. While the mismatch in variability between inlets and outlets is likely common across wetlands, the extent to which it affects estimates of phosphorus retention likely depends on the relationship between discharge and phosphorus concentrations, which could change with smaller wetlands, wetlands with less agricultural influence, or wetlands that are less driven by storm events. Broader studies suggest that the positive relationship between discharge and concentration in our case study is common, especially across catchments with anthropogenic disturbances, but a broader analysis on the range in both discharge-concentration relationships and differences in variability between the inlet and outlet in constructed wetlands is needed to understand how changing phosphorus sampling frequency affects estimates of retention by constructed wetlands (Ali et al., 2017; D'Amario et al., 2021; Moatar et al., 2017). However, our data highlights that a strong relationship is not needed to bias estimates of phosphorus retention. This is because storms represent a large portion of the water and phosphorus loading for wetlands, so even a few events (if the events are likely to have higher concentrations of phosphorus during storms) are enough to make large changes to nutrient retention calculations.

When comparing our case study with what we learned in the literature review, our data suggests that measurements of phosphorus retention in the literature may significantly underestimate the amount of phosphorus being retained by wetlands. Less than 30 % of studies from our literature review deliberately sample for surface water P concentrations during high flow events. Our case study is unlikely to represent the full range of potential biases when measuring wetland functionality; for instance, wetlands with negative discharge-concentration relationships would in fact have the opposite pattern of bias. An underestimate of phosphorus retention is most likely to be observed in wetlands that have precipitation-driven inflows and a positive discharge-concentration relationship, as is common in wetlands created in agricultural catchments (D'Amario et al., 2021).

4.3. Implications for monitoring and future research

While this review is focused on constructed and restored wetlands, the effects of sampling on bias can also apply to natural wetlands. Previous *meta*-analyses have found that wetlands with precipitation-driven inflows were most commonly sources of phosphorus, and mesocosms were rarely sources of phosphorus (Ury et al., 2023); however, our data suggests that those trends are likely influenced, if not caused, by biases in available data. This is further supported by studies showing that major rainfall events do not necessarily contribute to the release of phosphorus from wetlands (Dunne et al., 2015; Jordan et al., 2003; Tanner and Sukias, 2011). The variability in whether storm events release phosphorus may be more related to the potential for drying and rewetting of wetlands, which has been demonstrated to mobilize organic and labile phosphorus from the sediment (Kinsman-Costello et al., 2016; Kinsman-Costello et al., 2014). This can be further exacerbated by large stores of phosphorus within wetland soils, either from buildup over time or as a legacy of past land use (Nair et al., 2015; Sharpley et al., 2013; Wiegman et al., 2022). Understanding how phosphorus stored in sediments affects the ability of wetlands to retain phosphorus is a critical question for understanding the long-term function of wetlands.

Our synthesis shows that data on the long-term function of constructed/restored wetlands are limited, but that these wetlands have the potential to retain phosphorus for multiple decades. We found that long-term monitoring is needed to understand trajectories of wetlands and that for every year of monitoring one can accurately predict approximately one to two years into the future. Of the wetlands with more than ten years of data, only one was a phosphorus source, but this is likely influenced by a kind of "survivorship bias," in which monitoring is not continued for wetlands that are not retaining phosphorus, even though that data could be highly valuable to understand the effect of wetland sediments on nutrient release by constructed wetlands (Montgomery et al., 2021). None of the wetlands with estimates of long-term retention (greater than 10 years) documented in the available literature included event samples, which may further bias our understanding of the capacity for wetlands to retain phosphorus long-term. Our analysis of long-term data highlights the need for a greater variety in long-term wetland data, with only four studied wetlands treating sewage, and of the seven treating agricultural runoff, six were subsets of the Everglades Stormwater Treatment Areas (C. Pietro and Ivanoff, 2015; Jóźwiakowski et al., 2018; Kadlec, 2009; Ronkanen and Kløve, 2009; Vymazal, 2011). Our use of the broken window algorithm with these datasets highlights that short-term trends, which we show are the majority of wetland studies, are likely to create misleading wetland trajectories that are not representative of long-term trends in wetland phosphorus retention (Bahlai et al., 2021).

To our knowledge this is the first paper to analyze the effects of study time, and sampling frequency on the retention of phosphorus in constructed and restored wetlands. While we highlight the potential for bias across wetlands, along with a relevant mechanism (the retention of water by wetlands), the broader impacts of these sources of bias we have indicated require further analysis. For example, how resampling affects wetlands with different discharge-concentration relationships, and different residence times of water, is needed to better predict how different types of wetlands respond to different sampling frequencies. Similarly, an analysis on a broader variety of wetlands of different sizes and structures is needed to construct a predictive framework of what wetland features affect phosphorus retention.

5. Conclusions

Among the present literature, we identified two major data gaps and one major source of bias in our understanding of phosphorus retention by constructed and restored wetlands. We show that, especially in wetlands with precipitation-driven flow, there is likely a significant underestimation of nutrient retention when storm events are not accounted for, which is the case for most studies on wetland nutrient retention. We additionally highlight that studied wetlands are highly diverse and we are unable to explain variability in phosphorus retention using broad groupings. As wetlands are constructed for nutrient retention purposes, data on which attributes of wetlands influence the retention of nutrients is needed to improve nutrient mitigation efforts. To facilitate future efforts to understand mechanisms of wetland phosphorus retention, we recommend that data on flow regime, land-use history, water sources, water chemistry, sediment characteristics, and both vegetation abundance and species composition are reported in studies measuring wetland phosphorus retention, however further analysis is needed to develop a comprehensive list of important site metrics that influence phosphorus retention. Finally, we show that, while data on the long-term success of constructed wetlands for phosphorus retention is limited, that there are multiple examples of long-term nutrient storage in constructed wetlands and that short-term measurements are inadequate to represent long-term trends. Based on our findings, we recommend when monitoring constructed wetlands that, despite the very real difficulties of doing so, event sampling is considered and periods longer than the initial three years after construction are sampled. Not doing so risks inaccurate assessment of constructed wetlands and biases our understanding of phosphorus retention across wetlands in the literature.

Author contributions

KJA, BA, OFS, RMM, MPB, NB, JCG, SM, CB, DMC, and LK participated in the writing and editing of this manuscript; KA, BA, OFS, RMM, MPB, NB collected data for the literature analysis; KJA, BA, CB, DMC, LK contributed to data analyses for the literature analysis; KJA, BA, JCG, SM collected and analyzed data for the case study at Old Woman Creek; KJA, BA, OFS, RMM, MPB, CB, DMC, and LK conceptualized this manuscript.

CRediT authorship contribution statement

Kenneth J. Anderson: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Bishwodeep Adhikari:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Olivia F. Schloegel:** Writing – review & editing, Data curation, Conceptualization. **Raissa Marques Mendonca:** Writing – review & editing, Data curation, Conceptualization. **Michael P. Back:** Writing – review & editing, Data curation, Conceptualization. **Nicholas Brocato:** Writing – review & editing, Data curation. **Jacob A. Cianci-Gaskill:**

Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Steven E. McMurray:** Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Christie Bahlai:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **David M. Costello:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Lauren Kinsman-Costello:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This material is based upon work supported by the Great Lakes Restoration Initiative (GLRI) and funded under grant EPA-R5-GL00E03253. Development of the algorithm used in the trajectory analysis was supported by National Science Foundation Grants OAC 1838807 and DBI 2045721. The authors thank the members of the H2Ohio wetland monitoring team for advice and feedback on the development of this manuscript. The authors thank the H2Ohio GLRI Sensor Network Advisory Group for advice and feedback during the development of this manuscript: Nandita Basu, Whitney Beck, Jacob Berkowitz, Paul Gledhill, Heather Golden, Matt Pawlowski, Katie Quesnell, Craig Stow, Santina Wortman. The authors especially thank Nandita Basu for advising us to change course at the beginning of our literature analysis. The authors additionally thank Janice Kerns and Eric Saas from the Ohio Department of Natural Resources, United States for their advice and feedback during the development of this manuscript. The authors thank John Kominoski for the meta-analysis project in his advanced ecology class which taught Kenneth Anderson what to avoid during the literature review for this manuscript. The authors thank Robin Miller for editing, advice, and comments during the writing of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2024.112969.

Data availability

Data for the literature synthesis is included as a supplemental file to this manuscript. Data from the resampling case study at Old Woman Creek are available in the Environmental Data Initiative's Data Portal from McMurray et al (2024) at https://doi.org/10.6073/pasta/3f251395d82eafb33a93a9f3ebfb858c.

References

Ackerman, J.R., Peterson, E.W., Van der Hoven, S., Perry, W.L., 2015. Quantifying nutrient removal from groundwater seepage out of constructed wetlands receiving treated wastewater effluent. Environ. Earth Sci. 74, 1633–1645. https://doi.org/10.1007/s12665-015-4167-3.

Ali, G., Wilson, H., Elliott, J., Penner, A., Haque, A., Ross, C., Rabie, M., 2017. Phosphorus export dynamics and hydrobiogeochemical controls across gradients of scale, topography and human impact. Hydrol. Process. 31, 3130–3145. https://doi.org/10.1002/hyp.11258.

Anderson, D.M., Glibert, P.M., Burkholder, J.M., 2002. Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. Estuaries 25, 704–726. https://doi.org/10.1007/BF02804901.

Appling, A.P., Leon, M.C., McDowell, W.H., 2015. Reducing bias and quantifying uncertainty in watershed flux estimates: the R package loadflex. Ecosphere 6, 1–25. https://doi.org/10.1890/ES14-00517.1.

Ardón, M., Montanari, S., Morse, J.L., Doyle, M.W., Bernhardt, E.S., 2010. Phosphorus export from a restored wetland ecosystem in response to natural and experimental hydrologic fluctuations. J. Geophys. Res. Biogeo. 115. https://doi.org/10.1029/2009JG001169.

Aulenbach, B.T., 2013. Improving regression-model-based streamwater constituent load estimates derived from serially correlated data. J. Hydrol. 503, 55–66.

Aulenbach, B.T., Hooper, R.P., 2006. The composite method: an improved method for stream-water solute load estimation. Hydrol. Process. 20, 3029–3047. https://doi.org/10.1002/hyp.6147.

Aulenbach, B.T., Burns, D.A., Shanley, J.B., Yanai, R.D., Bae, K., Wild, A.D., Yang, Y., Yi, D., 2016. Approaches to stream solute load estimation for solutes with varying dynamics from five diverse small watersheds. Ecosphere 7, e01298. https://doi.org/10.1002/ecs2.1298.

Aziz, T., Van Cappellen, P., 2021. Economic valuation of suspended sediment and phosphorus filtration services by four different wetland types: a preliminary assessment for southern Ontario, Canada. Hydrol. Process. 35, e14442. https://doi.org/10.1002/hyp.14442.

Bahlai, C.A., White, E.R., Perrone, J.D., Cusser, S., Stack Whitney, K., 2021. The broken window: an algorithm for quantifying and characterizing misleading trajectories in ecological processes. Eco. Inform. 64, 101336. https://doi.org/10.1016/j.ecoinf.2021.101336.

Baird, R.B., Eaton, A.D., Rice, E.W. (Eds.), 2017. Standard Methods for the Examination of Water and Wastewater, 23rd ed. American Public Health Association, Washington, DC.

Birgand, F., Faucheux, C., Gruau, G., Augeard, B., Moatar, F., Bordenave, P., 2010. Uncertainties in assessing annual nitrate loads and concentration indicators: Part 1. Impact of sampling frequency and load estimation algorithms. Trans. ASABE 53, 437–446.

Braskerud, B.C., Lundekvam, H., Krogstad, T., 2000. The impact of hydraulic load and aggregation on sedimentation of soil particles in small constructed wetlands. J. Environ. Quality 29, 2013–2020. https://doi.org/10.2134/jeq2000.00472425002900060039x.

Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl. 8, 559–568.

Cheng, F.Y., Basu, N.B., 2017. Biogeochemical hotspots: Role of small water bodies in landscape nutrient processing. Water Resour. Res. 53, 5038–5056. https://doi.org/10.1002/2016WR020102.

Clifford, C.C., Heffernan, J.B., 2018. Artificial Aquatic Ecosystems. Water 10, 1096. https://doi.org/10.3390/w10081096.

Cowardin, L.M., 1979. Classification of wetlands and deepwater habitats of the United States. Fish and Wildlife Service, US Department of the Interior.

Cowardin, L.M., Golet, F.C., 1995. US Fish and Wildlife Service 1979 wetland classification: A review, in: Finlayson, C.M., Van Der Valk, A.G. (Eds.), Classification and Inventory of the World's Wetlands. Springer Netherlands, Dordrecht, pp. 139–152. https://doi.org/10.1007/978-94-011-0427-2_12.

Dell'Osbel, N., Colares, G.S., de Oliveira, G.A., de Souza, M.P., Barbosa, C.V., Machado, Ê.L., 2020. Bibliometric analysis of phosphorous removal through constructed wetlands. Water Air Soil Pollut. 231, 117. https://doi.org/10.1007/s11270-020-04513-1.

Djodjic, F., Geranmayeh, P., Markensten, H., 2020. Optimizing placement of constructed wetlands at landscape scale in order to reduce phosphorus losses. Ambio 49, 1797–1807. https://doi.org/10.1007/s13280-020-01349-1.

Dunne, E.J., Coveney, M.F., Hoge, V.R., Conrow, R., Naleway, R., Lowe, E.F., Battoe, L. E., Wang, Y., 2015. Phosphorus removal performance of a large-scale constructed treatment wetland receiving eutrophic lake water. Ecol. Eng. 79, 132–142. https://doi.org/10.1016/j.ecoleng.2015.02.003.

D'Amario, S.C., Wilson, H.F., Xenopoulos, M.A., 2021. Concentration-discharge relationships derived from a larger regional dataset as a tool for watershed management. Ecol. Appl. 31, e02447. https://doi.org/10.1002/eap.2447.

Gerbeaux, P., Finlayson, C.M., Van Dam, A., 2018. Wetland classification: overview. Structure and function, management and methods, The Wetland book I, pp. 1461–1468.

Ghosh, D., Gopal, B., 2010. Effect of hydraulic retention time on the treatment of secondary effluent in a subsurface flow constructed wetland. Ecol. Eng. 36, 1044–1051. https://doi.org/10.1016/j.ecoleng.2010.04.017.

Golden, H.E., Rajib, A., Lane, C.R., Christensen, J.R., Wu, Q., Mengistu, S., 2019. Non-floodplain wetlands affect watershed nutrient dynamics: a critical review. Environ. Sci. Tech. 53, 7203–7214. https://doi.org/10.1021/acs.est.8b07270.

Gordon, B.A., Dorothy, O., Lenhart, C.F., 2020. Nutrient retention in ecologically functional floodplains: a review. Water 12, 2762. https://doi.org/10.3390/w12102762.

Hambäck, P.A., Dawson, L., Geranmayeh, P., Jarsjö, J., Kačergytė, I., Peacock, M., Collentine, D., Destouni, G., Futter, M., Hugelius, G., Hedman, S., Jonsson, S., Klatt, B.K., Lindström, A., Nilsson, J.E., Pärt, T., Schneider, L.D., Strand, J.A., Urrutia-Cordero, P., Åhlén, D., Åhlén, I., Blicharska, M., 2023. Tradeoffs and synergies in wetland multifunctionality: a scaling issue. Sci. Total Environ. 862, 160746. https://doi.org/10.1016/j.scitotenv.2022.160746.

Heisler, J., Glibert, P.M., Burkholder, J.M., Anderson, D.M., Cochlan, W., Dennison, W. C., Dortch, Q., Gobler, C.J., Heil, C.A., Humphries, E., Lewitus, A., Magnien, R., Marshall, H.G., Sellner, K., Stockwell, D.A., Stoecker, D.K., Suddleson, M., 2008. Eutrophication and harmful algal blooms: a scientific consensus. Harmful Algae, HABs and Eutrophication 8, 3–13. https://doi.org/10.1016/j.hal.2008.08.006.

Johnes, P.J., 2007. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. J. Hydrol. 332, 241–258. https://doi.org/10.1016/j.jhydrol.2006.07.006.

Johnson, L., 2018. Heidelberg Tributary Loading Program: Old Woman Creek Special Study. National Center for Water Quality Research.

Jordan, T.E., Whigham, D.F., Hofmockel, K.H., Pittek, M.A., 2003. Nutrient and sediment removal by a restored wetland receiving agricultural runoff. J. Environ. Qual. 32, 1534–1547. https://doi.org/10.2134/jeq2003.1534.

Jóźwiakowski, K., Bugajski, P., Kurek, K., de Fátima Nunes de Carvalho, M., Almeida, M. A.A., Siwiec, T., Borowski, G., Czekała, W., Dach, J., Gajewska, M., 2018. The efficiency and technological reliability of biogenic compounds removal during long-term operation of a one-stage subsurface horizontal flow constructed wetland. Sep. Purif. Technol. 202, 216–226. https://doi.org/10.1016/j.seppur.2018.03.058.

Kadlec, R.H., 2009. Wastewater treatment at the Houghton Lake wetland: Hydrology and water quality. Ecological Engineering, the Houghton Lake Wetland Treatment Project 35, 1287–1311. https://doi.org/10.1016/j.ecoleng.2008.10.001.

Kamrath, B., Yuan, Y., Manning, N., Johnson, L., 2023. Influence of sampling frequency and estimation method on phosphorus load uncertainty in the Western Lake Erie Basin, Ohio, USA. J. Hydrol. 617, 128906. https://doi.org/10.1016/j.jhydrol.2022.128906.

Kelly, P.T., Vanni, M.J., Renwick, W.H., 2018. Assessing uncertainty in annual nitrogen, phosphorus, and suspended sediment load estimates in three agricultural streams using a 21-year dataset. Environ. Monit. Assess. 190, 91. https://doi.org/10.1007/s10661-018-6470-4.

Kiedrzyńska, E., Wagner, I., Zalewski, M., 2008. Quantification of phosphorus retention efficiency by floodplain vegetation and a management strategy for a eutrophic reservoir restoration. Ecological Engineering 33, 15–25. https://doi.org/10.1016/j.ecoleng.2007.10.010.

Kinsman-Costello, L.E., O'Brien, J., Hamilton, S.K., 2014. Re-flooding a historically drained wetland leads to rapid sediment phosphorus release. Ecosystems 17, 641–656. https://doi.org/10.1007/s10021-014-9748-6.

Kinsman-Costello, L.E., Hamilton, S.K., O'Brien, J.M., Lennon, J.T., 2016. Phosphorus release from the drying and reflooding of diverse shallow sediments. Biogeochemistry 130, 159–176. https://doi.org/10.1007/s10533-016-0250-4.

Krieger, K.A. 2003. Effectiveness of a coastal wetland in reducing pollution of a Laurentian Great Lake: Hydrology, sediment, and nutrients. Wetlands 23:778–791. doi: 10.1672/0277-5212(2003)023[0778:EOACWI]2.0.CO;2.

Krijthe, J., 2014. Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut Implementation. R Package Version 17. https://github.com/jkrijthe/Rtsne.

Krzeminska, D., Blankenberg, A.-G.B., Bechmann, M., 2023. The effectiveness of sediment and phosphorus removal by a small constructed wetland in Norway: 18 years of monitoring and perspectives for the future. CATENA 223, 106962. https://doi.org/10.1016/j.catena.2023.106962.

Land, M., Granéli, W., Grimvall, A., Hoffmann, C.C., Mitsch, W.J., Tonderski, K.S., Verhoeven, J.T.A., 2016. How effective are created or restored freshwater wetlands for nitrogen and phosphorus removal? A systematic review. Environ. Evidence 5, 9. https://doi.org/10.1186/s13750-016-0060-0.

Leibowitz, S.G., Hill, R.A., Creed, I.F., Compton, J.E., Golden, H.E., Weber, M.H., Rains, M.C., Jones, C.E., Lee, E.H., Christensen, J.R., Bellmore, R.A., Lane, C.R., 2023. National hydrologic connectivity classification links wetlands with stream water quality. Nat. Water 1, 370–380. https://doi.org/10.1038/s44221-023-00057-w.

Li, S., Elliott, J.A., Tiessen, K.H.D., Yarotski, J., Lobb, D.A., Flaten, D.N., 2011. The Effects of Multiple Beneficial Management Practices on Hydrology and Nutrient Losses in a Small Watershed in the Canadian Prairies. J. Environ. Qual. 40, 1627–1642. https://doi.org/10.2134/jeq2011.0054.

Li, D., Chu, Z., Li, P., Xu, W., Wang, E., Jin, C., Zheng, B., 2023. Impacts of landscape spatial configuration of integrated multi-pond constructed wetlands in a basin on the treatment of non-point source pollution. J. Clean. Prod. 383, 135389. https://doi.org/10.1016/j.jclepro.2022.135389.

Lindstrom, S.M., White, J.R., 2011. Reducing phosphorus flux from organic soils in surface flow treatment wetlands. Chemosphere 85, 625–629. https://doi.org/10.1016/j.chemosphere.2011.06.109.

Locke, D.H., Grove, J.M., 2016. Doing the Hard Work Where it's Easiest? Examining the Relationships Between Urban Greening Programs and Social and Ecological Characteristics. Appl. Spatial Analysis 9, 77–96. https://doi.org/10.1007/s12061-014-9131-1.

Maechler, M., Rousseuw, P., Struyf, A., Hubert, M., Hornik, K., 2023. cluster: Cluster Analysis Basics and Extensions. R Package Version 2 (1), 6. https://CRAN.R-project.org/package=cluster.

Marton, J.M., Creed, I.F., Lewis, D.B., Lane, C.R., Basu, N.B., Cohen, M.J., Craft, C.B., 2015. Geographically isolated wetlands are important biogeochemical reactors on the landscape. Bioscience 65, 408–418.

Mitsch, W.J., 1992. Landscape design and the role of created, restored, and natural riparian wetlands in controlling nonpoint source pollution. Ecol. Eng. 1, 27–47. https://doi.org/10.1016/0925-8574(92)90024-V.

Mitsch, W.J., Reeder, B.C., 1992. Nutrient and hydrologic budgets of a great lakes coastal freshwater wetland during a drought year. Wetlands Ecology and Management 1, 211–222. https://doi.org/10.1007/BF00244926.

Moatar, F., Abbott, B.W., Minaudo, C., Curie, F., Pinay, G., 2017. Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. Water Resour. Res. 53, 1270–1287. https://doi.org/10.1002/2016WR019635.

Montgomery, J.A., Eames, J.M., Klimas, C., 2021. A 16-year investigation of legacy phosphorus discharge from Prairie Wolf Slough: a wetland restored on a former farmed field. Restor. Ecol. 29, e13340. https://doi.org/10.1111/rec.13340.

Nair, V.D., Clark, M.W., Reddy, K.R., 2015. Evaluation of Legacy Phosphorus Storage and Release from Wetland Soils. J. Environ. Qual. 44, 1956–1964. https://doi.org/10.2134/jeq2015.03.0154.

Persson, J., Somes, N.L.G., Wong, T.H.F., 1999. Hydraulics efficiency of constructed wetlands and ponds. Water Sci. Technol. 40, 291–300.

Pietro, C.K., Ivanoff, D., 2015. Comparison of long-term phosphorus removal performance of two large-scale constructed wetlands in South Florida, U.S.A. Ecol. Eng. 79, 143–157. https://doi.org/10.1016/j.ecoleng.2014.12.013.

Powers, S.M., Julian, J.P., Doyle, M.W., Stanley, E.H., 2013. Retention and transport of nutrients in a mature agricultural impoundment. J. Geophys. Res. Biogeo. 118, 91–103. https://doi.org/10.1029/2012JG002148.

R Core Team, 2024. R: A language and environment for statistical computing.

Reddy, K.R., Kadlec, R.H., Flaig, E., Gale, P.M., 1999. Phosphorus retention in streams and wetlands: a review. Crit. Rev. Environ. Sci. Technol. 29, 83–146.

Reddy, K.R., DeLaune, R.D., Inglett, P.W., 2022. Biogeochemistry of wetlands: science and applications. CRC Press.

Richards, R.P., Holloway, J., 1987. Monte Carlo studies of sampling strategies for estimating tributary loads. Water Resour. Res. 23, 1939–1948. https://doi.org/10.1029/WR023i010p01939.

Robotham, J., Old, G., Rameshwaran, P., Sear, D., Trill, E., Bishop, J., Gasca-Tucker, D., Old, J., McKnight, D., 2023. Nature-based solutions enhance sediment and nutrient storage in an agricultural lowland catchment. Earth Surf. Proc. Land. 48, 243–258. https://doi.org/10.1002/esp.5483.

Ronkanen, A.-K., Kløve, B., 2009. Long-term phosphorus and nitrogen removal processes and preferential flow paths in Northern constructed peatlands. Ecol. Eng. 35, 843–855. https://doi.org/10.1016/j.ecoleng.2008.12.007.

Schmadel, N.M., Harvey, J.W., Schwarz, G.E., Alexander, R.B., Gomez-Velez, J.D., Scott, D., Ator, S.W., 2019. Small Ponds in Headwater Catchments Are a Dominant Influence on Regional Nutrient and Sediment Budgets. Geophys. Res. Lett. 46, 9669–9677. https://doi.org/10.1029/2019GL083937.

Sharpley, A., Jarvie, H.P., Buda, A., May, L., Spears, B., Kleinman, P., 2013. Phosphorus Legacy: Overcoming the Effects of Past Management Practices to Mitigate Future Water Quality Impairment. J. Environ. Qual. 42, 1308–1326. https://doi.org/10.2134/jeq2013.03.0098.

Shen, W., Li, S., Basu, N.B., Ury, E.A., Jing, Q., Zhang, L., 2023. Size and temperature drive nutrient retention potential across water bodies in China. Water Res. 239, 120054. https://doi.org/10.1016/j.watres.2023.120054.

Smith, V.H., 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. Environ. Sci. Pollut. Res. 10, 126–139. https://doi.org/10.1065/espr2002.12.142.

Stammel, B., Amtmann, M., Gelhaus, M., Cyffka, B., 2018. Change of regulating ecosystem services in the Danube floodplain over the past 150 years induced by land use change and human infrastructure. DIE ERDE – J. Geogr. Soc. Berlin 149, 145–156.

Steinman, A.D., Ogdahl, M.E., 2011. Does converting agricultural fields to wetlands retain or release P? J. N. Am. Benthol. Soc. 30, 820–830. https://doi.org/10.1899/10-106.1.

Steinman, A.D., Hassett, M., Oudsema, M., 2018. Effectiveness of Best Management Practices to Reduce Phosphorus Loading to a Highly Eutrophic Lake. Int. J. Environ. Res. Public Health 15, 2111. https://doi.org/10.3390/ijerph15102111.

Tanner, C.C., Sukias, J.P.S., 2011. Multiyear nutrient removal performance of three constructed wetlands intercepting tile drain flows from grazed pastures. J. Environ. Qual. 40, 620–633. https://doi.org/10.2134/jeq2009.0470.

Tiner, R.W., 2015. Classification of wetland types for mapping and large-scale inventories. In: Remote Sensing of Wetlands: Applications and Advances. CRC Press, pp. 19–42.

Tomscha, S.A., Bentley, S., Platzer, E., Jackson, B., de Roiste, M., Hartley, S., Norton, K., Deslippe, J.R., 2021. Multiple methods confirm wetland restoration improves ecosystem services. Ecosyst. People 17, 25–40. https://doi.org/10.1080/26395916.2020.1863266.

Turner, V.K., Jarden, K., Jefferson, A., 2016. Resident perspectives on green infrastructure in an experimental suburban stormwater management program. Cities and the Environment (CATE) 9.

Ury, E.A., Arrumugam, P., Herbert, E.R., Badiou, P., Page, B., Basu, N.B., 2023. Source or sink? Meta-analysis reveals diverging controls of phosphorus retention and release in restored and constructed wetlands. Environ. Res. Lett. 18, 083002.

Uusi-Kämppä, J., Braskerud, B., Jansson, H., Syversen, N., Uusitalo, R., 2000. Buffer zones and constructed wetlands as filters for agricultural phosphorus. J. Env. Quality 29, 151–158. https://doi.org/10.2134/jeq2000.00472425002900010019x.

Vymazal, J., 2007. Removal of nutrients in various types of constructed wetlands. Science of the Total Environment, Contaminants in Natural and Constructed Wetlands: Pollutant Dynamics and Control 380, 48–65. https://doi.org/10.1016/j.scitotenv.2006.09.014.

Vymazal, J., 2011. Long-term performance of constructed wetlands with horizontal sub-surface flow: Ten case studies from the Czech Republic. Ecological Engineering, Special Issue: Enhancing Ecosystem Services on the Landscape with Created, Constructed and Restored Wetlands 37, 54–63. https://doi.org/10.1016/j.ecoleng.2009.11.028.

Walling, D., Webb, W., 1981. The reliability of suspended sediment load data.

Walton, C.R., Zak, D., Audet, J., Petersen, R.J., Lange, J., Oehmke, C., Wichtmann, W., Kreyling, J., Grygoruk, M., Jabłońska, E., Kotowski, W., Wiśniewska, M.M., Ziegler, R., Hoffmann, C.C., 2020. Wetland buffer zones for nitrogen and phosphorus retention: Impacts of soil type, hydrology and vegetation. Sci. Total Environ. 727, 138709. https://doi.org/10.1016/j.scitotenv.2020.138709.

Weisner, S.E.B., Johannesson, K., Thiere, G., Svengren, H., Ehde, P.M., Tonderski, K.S., 2016. National large-scale wetland creation in agricultural areas—potential versus realized effects on nutrient transports. Water 8, 544. https://doi.org/10.3390/w8110544.

White, J.R., Reddy, K. Ramesh, Moustafa, M.Z. 2004. Influence of hydrologic regime and vegetation on phosphorus retention in Everglades stormwater treatment area wetlands. Hydrological Processes 18:343–355. doi: 10.1002/hyp.1379.
Wickham, H., Wickham, M.H., 2017. Package 'tidyr.' Easily Tidy Data with 'spread' and 'gather (.)' Functions.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. J. Open Source Softw. 4, 1686. https://doi.org/10.21105/joss.01686.
Wickham, H., Francois, R., Henry, L., Muller, K., 2023. dplyr: A Grammar of Data Manipulation. R package version 1.1.4. https://github.com/tidyverse/dplyr, https://dplyr.tidyverse.org.
Widney, S., Kanabrocki Klein, A., Ehman, J., Hackney, C., Craft, C., 2018. The value of wetlands for water quality improvement: an example from the St. Johns River watershed, Florida. Wetlands Ecol Manage 26, 265–276. https://doi.org/10.1007/s11273-017-9569-4.
Wiegman, A.R.H., Myers, G.H., Augustin, I.C., Kubow, M.L., Fein-Cole, M.J., Perillo, V.L., Ross, D.S., Diehl, R.M., Underwood, K.L., Bowden, W.B., Roy, E.D., 2022. Potential for soil legacy phosphorus release from restored riparian wetlands within an agricultural landscape. Biogeochemistry 161, 137–156. https://doi.org/10.1007/s10533-022-00972-2.
Yan, W., Yin, C., Tang, H., 1998. Nutrient retention by multipond systems: mechanisms for the control of nonpoint source pollution. J. Environ. Quality 27, 1009–1017. https://doi.org/10.2134/jeq1998.00472425002700050005x.
Zhang, Z., Wang, Z., Wang, Y., Chen, X., Wang, H., Xu, X., XianYong, L., Czapar, G.F., 2011. Properties of phosphorus retention in sediments under different hydrological regimes: A laboratory-scale simulation study. Journal of Hydrology 404, 109–116. https://doi.org/10.1016/j.jhydrol.2010.06.018.
Zhu, Q., Schmidt, J.P., Bryant, R.B., 2012. Hot moments and hot spots of nutrient losses from a mixed land use watershed. J. Hydrol. 414–415, 393–404. https://doi.org/10.1016/j.jhydrol.2011.11.011.