Sediment and Nutrient Retention in Ponds on an Agricultural Stream: Evaluating Effectiveness for Diffuse Pollution Mitigation

John Robotham 1,2,*, Gareth Old 1, Ponnambalam Rameshwaran 1, David Sear 2, David Gasca-Tucker 3, James Bishop 4,4, Joanne Old 5 and David McKnight 5

1 UK Centre for Ecology and Hydrology, Wallingford, Oxfordshire OX10 8BB, UK; gho@ceh.ac.uk (G.O.); ponr@ceh.ac.uk (P.R.); jambis@ceh.ac.uk (J.B.)
2 School of Geography and Environmental Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK; d.sear@soton.ac.uk
3 Atkins, One St Aldates, Oxford OX1 1DE, UK; David.GascaTucker@atkinsglobal.com
4 School of Archaeology, Geography and Environmental Science, University of Reading, Whiteknights, Reading RG6 6UR, UK
5 Environment Agency, Wallingford, Oxfordshire OX10 8BD, UK; joanne.old@environment-agency.gov.uk (J.O.); david.mcknight@environment-agency.gov.uk (D.M.)
* Correspondence: johrob@ceh.ac.uk

Abstract: The creation of ponds and wetlands has the potential to alleviate stream water quality impairment in catchments affected by diffuse agricultural pollution. Understanding the hydrological and biogeochemical functioning of these features is important in determining their effectiveness at mitigating pollution. This study investigated sediment and nutrient retention in three connected (on-line) ponds on a lowland headwater stream by sampling inflowing and outflowing concentrations during base and storm flows. Sediment trapping devices were used to quantify sediment and phosphorus accumulations within ponds over approximately monthly periods. The organic matter content and particle size composition of accumulated sediment were also measured. The ponds retained dissolved nitrate, soluble reactive phosphorus and suspended solids during baseflows. During small to moderate storm events, some ponds were able to reduce peak concentrations and loads of suspended solids and phosphorus; however, during large magnitude events, resuspension of deposited sediment resulted in net loss. Ponds filtered out larger particles most effectively. Between August 2019 and March 2020, the ponds accumulated 0.306 t ha$^{-1}$ sediment from the 30 ha contributing area. During this period, total sediment accumulations in ponds were estimated to equal 7.6% of the suspended flux leaving the 340 ha catchment downstream. This study demonstrates the complexity of pollutant retention dynamics in on-line ponds and highlights how their effectiveness can be influenced by the timing and magnitude of events.

Keywords: water quality; natural flood management; suspended sediment; phosphorus; nitrate; organic matter; on-line ponds; constructed wetlands; particle size

1. Introduction

Intensively farmed landscapes can contribute significantly to the degradation of the water environment globally [1–3]. In many European countries, agricultural intensification has increased the risk of waterbodies failing to meet the EU Water Framework Directive (WFD) objective of ‘good ecological status’ (2000/60/EC) [4–6]. For streams and rivers in England and Wales, diffuse sources of pollution from agriculture present one of the biggest threats to WFD failure, with key concerns being elevated nutrient concentrations, oxygen depletion and the smothering of instream habitats by fine sediment [7]. One of the main delivery mechanisms for diffuse pollution is soil erosion caused by surface run-off which is then exacerbated when arable fields are left bare or subject to soil compaction [6,8]. Diffuse
agricultural pollution in the UK is estimated to have an annual cost of £238 million, resulting from reduced water quality and associated treatments costs [9]. These costs consider the wide-ranging negative impacts on biodiversity, ecosystem services, landscape value, rural public access and enjoyment, water and air quality, and natural resources. Additionally, soil erosion can lead to increased sedimentation in watercourses and a reduction in channel capacity, thereby increasing flood risk which is already an increasing concern as a result of climate change [10]. Given the substantial legislative, ecological and economic implications of diffuse agricultural pollution, there is a growing interest in the cost-effective delivery of measures to mitigate its negative effects. For example, the £6.5 million government-funded Demonstration Test Catchments (DTC) project (2009–2014) focused on four agriculturally representative catchments in England with the aim of generating evidence on how diffuse pollution can be controlled to improve water quality at multiple spatial scales [11]. Mitigation measures are wide-ranging and require robust data to be able to evaluate and compare their effectiveness to enable cost-effective implementation for future land management [12].

One commonly adopted mitigation measure is the creation of pond features on or adjacent to watercourses to intercept pollutants such as sediment, nutrients and pesticides during their transport instream or along overland pathways. These types of features have varying designs and are often referred to in the literature under different terms such as constructed wetlands, retention ponds, in-line/on-line ponds, settling ponds, run-off attenuation features and rural sustainable drainage systems (RSuDS) [12–16]. Differences in design mean they can be suited to specific agricultural, rural or urban contexts, but generally they aim to achieve the same purpose of improving water quality. These types of features have been implemented and studied across countries worldwide, notably in northern and western Europe. In France and England, pond systems have been used to treat motorway run-off, removing heavy metals such as cadmium by up to 100% in some cases [17,18]. Since as early as the 1960s, countries including Denmark and Germany have used constructed wetlands to treat domestic wastewater to reduce biochemical oxygen demand (BOD and also remove nitrogen (N) and phosphorus (P), typically by 30–50% [19,20]. Wetland features have been used in countries such as Norway, Finland and Estonia to treat nonpoint (diffuse) sources of agricultural pollution by removing suspended solids and various nutrient species, primarily total P [21–23]. A systematic review of created wetlands (mostly in North America and Europe) found that on average they significantly reduce P and N transport from wastewater, urban and agricultural run-off, with median removal rates of 1.2 and 93 g m$^{-2}$ year$^{-1}$, respectively [24]. It was concluded that further research is needed on the effects of hydrological pulses on wetlands, as it was found that wetlands with nutrient loading rates driven by rainfall had significantly lower P removal efficiencies than wetlands with controlled loading rates.

In addition to diffuse pollution mitigation, such features have the potential to provide co-benefits (ecosystem services) including carbon storage, flood risk reduction, low flow and drought resilience, habitat provision and aesthetic quality [12,25–31]. Evidence on the provision of some of these services has been reported with varying degrees of effectiveness for different benefits and for different designs, climates, soils and catchment characteristics [12,21,32,33]. However, it is often still assumed that these pond features continuously deliver co-benefits without consideration into how their effectiveness may change over time as a consequence of the magnitude of storm events they experience and how they are managed or maintained.

Evidence on the site-specific constraints and limitations of interventions is important to develop guidance that enables targeted mitigation for achieving maximal societal and ecosystem benefit. Contrasting catchments have different water quality issues of focus, making evidence of intervention effectiveness for different pollutants useful for catchment management plans. For example, in hydrologically ‘flashy’ catchments dominated by slowly permeable clay soils, where surface run-off generation is high, the key concern is typically the resulting high loads of fine suspended sediment and particle-bound P entering watercourses [34]. Despite their small size and discharge, agricultural drainage ditches and
streams in headwaters often form highly connected dense networks that can cumulatively convey large quantities of water, sediment and nutrients rapidly downstream [35]. This makes agricultural headwater streams an appropriate target area for both pollution and flood risk mitigation opportunities. The study catchment discussed in this paper has been managed to exploit these mitigation opportunities in recent years, as part of both a Natural Flood Management (NFM) scheme and a scheme to tackle diffuse agricultural pollution and address current failure to meet WFD targets. The ponds in this study were implemented primarily to help mitigate diffuse P and fine sediment pollution to manage water quality, alongside providing co-benefits for habitat creation and biodiversity.

Past studies examining the trapping efficiencies of constructed wetlands have suggested that greater retention of sediment can occur with increases in the run-off received per wetland surface area (hydraulic load), and that even small-scale features can be suitable for retaining fine clay particles [36,37]. More recent efforts studying multiple events have raised questions surrounding a lack of consistency in trapping efficiency, with instances of net loss of material being reported [13,14]. More evidence on the controls on trapping efficiencies for different sediment and nutrient fractions under a broad range of hydrological conditions will help to clarify this issue. Previous guidance on the design of constructed wetlands recommended an optimal width-to-length ratio of 1:4; however, in practice this can be difficult to achieve on farms with limited space [38]. Further evidence on the trapping efficiencies of different styles of pond and wetland (including those with suboptimal designs) can provide valuable information for developing sustainable designs for their future use in a variety of landscape contexts, including NFM schemes.

In the UK, ponds and wetlands are currently not widely used as interventions for mitigating the impacts of diffuse agricultural pollution, even more so in the case of on-line features that connect to existing streams. This paper aims to provide evidence on the effectiveness of small on-line ponds with ~1:1 aspect ratios to function as diffuse pollution interventions in the context of a lowland arable catchment in the UK. We quantify key water quality benefits derived from the pond system under both baseflows and stormflows, with a particular focus on its effectiveness to trap sediment and phosphorus over multiple hydrological events. It was hypothesised that the ponds would be most effective at trapping sediment, reducing both suspended sediment and phosphorus concentrations and loads downstream of each pond. This study also quantifies the accumulation of sediment within ponds since their construction in order to evaluate their sustainability and maintenance requirements in the long-term.

2. Materials and Methods

2.1. Study Site

The study site is located in the predominantly arable 16.3 km² Littlestock Brook subcatchment that lies within the Evenlode catchment (430 km²) in the upper reaches of the Thames basin in southern England, United Kingdom (Figure 1). Lithology within the Evenlode catchment is dominated by the Great Oolite Group, consisting of mudstone and fine-grained limestone (Figure 1b). The Littlestock Brook subcatchment on the western side of the Evenlode is mostly underlain by the Lias Group, consisting of clays, mudstones and limestones.

The Littlestock Brook subcatchment is being intensively monitored as part of a NFM pilot scheme that is being delivered over five years (2016–2021) to reduce flood risk in the village of Milton-under-Wychwood. A full description of the NFM project, its delivery, and interventions, is given by Old et al. (2019) [39].
The ponds are situated in a first-order headwater tributary of the Littlestock Brook close to its source which rises from a limestone geology overlain by a shallow lime-rich soil that transitions down the valley into seasonally wet, slowly permeable clay soils. The average saturated hydraulic conductivity of topsoil in the catchment is approximately 50 cm day\(^{-1}\) [40]. The area experiences a temperate maritime climate, with an average annual minimum temperature of 5.7 °C and maximum of 13.1 °C and receives an average annual rainfall of 765 mm [41]. The ponds drain an area of 0.3 km\(^2\) and only occupy <0.2% of this catchment. Over 75% of this catchment area is underlain by a highly productive fissured aquifer. Both the rest of this area and the downstream catchment (3.4 km\(^2\) area outlined in white; Figure 1c) are underlain by rocks with essentially no groundwater. This downstream catchment forms part of the wider high temporal resolution hydrological monitoring network for the NFM scheme, gauging flows and fluxes of waters/suspended matter leaving the NFM-impacted catchment outlet (detailed in Section 2.2). The ponds are situated in a steeper part of the catchment where run-off risk was identified as being high and overland flows had previously been observed.

The field containing the ponds was partly taken out of agricultural production for construction of the ponds and channel in February 2018 and was then surrounded by an area of mixed deciduous tree species planted in early 2019. The three ponds were dug out and the excavated soil used to form earth banks covering the outflows, which

---

**Figure 1.** Locations of (a) the Evenlode catchment (green) within the Thames basin (blue); (b) the Downstream Catchment (outlined in red) within the Evenlode catchment and its geology; (c) the field site of the on-line ponds (outlined in red) within the Downstream Catchment; (d) the on-line ponds and monitoring sites (labelled A-D from upstream to downstream) and (e) sediment traps within the Central Pond. The white field contains improved grassland used as a horse paddock. Contours (grey) display elevation (metres AODN). Crop data are from UKCEH Land Cover® plus: Crops (2018).
are comprised of layers of locally-sourced limestone that slowly allow flow through into the following stream reach. The ponds are teardrop-shaped and vary in size with the Upstream Pond having an estimated surface area and total capacity of 145 m² and 70 m³, the Central Pond 126 m² and 90 m³, and the Downstream Pond 156 m² and 95 m³ respectively (Figure 2) (estimates derived from pond cross-section surveys described in Section 2.3). The Upstream/Downstream ponds have width-to-length ratios of ~1:1, and ~1:1.5 for the Central Pond. The earth banks and ponds were not seeded and left to colonise naturally.

![Figure 2. The Central Pond during (a) low flow conditions in July 2019; and (b) at capacity during a storm event on 15 February 2020. Both photos are taken facing upstream.](image)

The newly dug stream channel is small and shallow with an average width of ~0.5 m and depth of ~0.15 m, and a gradient of 2.5%. This contrasts to adjacent streams running along the field margins, which both have deeply incised channels. The channel planform is relatively straight, however colonisation and growth of graminoid vegetation within the channel has started to form multiple channels in places. In June 2018, a water level sensor (Rugged TROLL 100, In-Situ; Redditch, UK) was deployed in the Central Pond to measure changes in water depth and temperature at 5-min intervals. Atmospheric pressure from a nearby (<1 km) sensor was also logged and used to compensate for changes in barometric pressure. Sensor water depth was then calibrated against the observed depth on the stage board (pictured in Figure 2a) measured during site visits. The sensors were set to log at 5-min intervals in order to capture rapid changes in pond water level due to the ‘flashy’ nature of the catchment during rainfall events. In February 2019, a tipping bucket rain gauge (Casella; Sycamore, IL, USA) was installed in an adjacent clearing to measure rainfall at 2-min intervals. A storage rain gauge was also installed at the same location to aid quality control of the tipping bucket gauge. During site visits, stored rainwater was emptied into a graduated cylinder and the volume checked against the tipping bucket rainfall total for the same period to ensure measurements were within a 5% tolerance range.

### 2.2. Water Quality Sampling and Analysis

To monitor water quality under near baseflow conditions, water samples from the on-line pond system’s inlet and outlet were collected during field visits every 2–4 weeks. One unfiltered 60 mL sample was taken for total phosphorus (TP), and two 60 mL samples were immediately filtered through a 0.45 μm cellulose nitrate membrane (Whatman™ WCN grade; Maidstone, UK) for analysis of total dissolved phosphorus (TDP), soluble
reactive phosphorus (SRP), and dissolved major ions (NO$_2^-$, NO$_3^-$, NH$_4^+$, F$^-$, Cl$^-$ and SO$_4^{2-}$). Particulate phosphorus (PP) was taken to be the difference between TP and TDP. Approximately 500 mL was sampled using a US DH-48 sampler for determination of suspended sediment concentration (SSC) and volatile solids concentration (VSC) (as a proxy for organic matter). Water chemistry samples were refrigerated at 4 °C upon return from the field until they were analysed following Wallingford Nutrient Chemistry Laboratories procedures described in detail by Bowes et al. (2018) [42]. SSC was determined gravimetrically by filtering known volumes of water samples through pre-ashed, dried and weighed Whatman™ GF/C™ filter papers, which were then oven dried at 105 °C for at least 2 h. Filter papers were then reweighed after cooling in a desiccator for 30 min. VSC was then determined through loss-on-ignition (LOI) by igniting filter papers in a muffle furnace (AAF 1100, Carbolite Gero; Hope, Derbyshire, UK) at 500 °C for 30 min before being cooled and reweighed [43].

For monitoring storm events, automatic samplers (Sigma SD900, Hach; Loveland, CO, USA) were deployed at four locations along the stream to sample water flowing into and out of each pond (Figure 1c). Triggering of samplers was determined based on the rainfall forecast in order to capture samples approximately representative of the event. Grab samples of run-off were taken from contributing overland flow pathways. Samples were refrigerated upon return to the laboratory, and 60 mL subsamples were taken as soon as possible for chemical determinands of interest. To ensure representative subsampling, samples were thoroughly mixed before immediately taking an aliquot using a syringe. The remaining sample was used to determine SSC and VSC. Discharge was estimated at the ponds’ outflows in higher flows using an Electromagnetic Current Meter (Valeport; Totnes, UK) and the velocity-area method [44], and also under low flows using a conductivity sensor (EXO1, YSI; Yellow Springs, OH, USA) and the salt dilution method [45]. During storm events, run-off frequently overwhelmed the small stream channel and rendered it unsuitable for accurate flow measurement or development of a reliable stage-discharge relationship. Instead, water flowing through the ponds was estimated as a catchment area-weighted proportion of the discharge measured at a more stable gauging site (Downstream Catchment Outlet; Figure 1c). In order to represent timings of storm hydrographs more realistically, the estimated discharge was shifted back in time by applying a linear regression ($R^2 = 0.51$) between peak discharge and the time difference between peak stage in the Central Pond and at the Downstream Catchment Outlet. It was assumed that at a given time, discharge was equal at both pond inflows and outflows.

The fluxes of total suspended sediment, silt and clay, and TP were also calculated at the Downstream Catchment Outlet site using discharge and SSC/TP data at 5-min intervals. Discharge was estimated using a stage-discharge rating curve with flow measurements taken using the methods described above, with measured discharges ranging from 6 to 587 L s$^{-1}$ ($n = 15$). Turbidity was monitored using an in-stream sensor (DTS-12, FTS; Victoria, Canada) and then calibrated against SSC and TP samples ($R^2 = 0.99$, $n = 95$; $R^2 = 0.79$, $n = 372$) taken under a range of flows (sampled using the methods described above) to give estimated timeseries of SSC and TP. Turbidity data covered >99% of the monitoring period. Suspect datapoints were removed and the gaps filled by linear interpolation for periods of <12 h if no storm events took place during the missing period. Fluxes were calculated by integrating SSC/TP instantaneous load timeseries for the monitoring period. Suspended sediment particle size distributions were also sampled ($n = 9$) during two high flow/SSC events (measured using laser diffraction as described in Section 2.3). These event particle size distributions were assumed to be representative of the stream’s suspended load as storm events contribute the majority of the total sediment flux. The proportions of particles <$63 \mu$m in diameter in the samples were averaged and combined to estimate the flux of silt and clay leaving the catchment.
2.3. Pond Sediment Sampling and Analysis

Sediment traps were deployed in each pond to quantify sediment, organic matter, and P accumulation, and determine particle size distribution. Traps were assembled from circular plastic saucers (19 cm in diameter, 4 cm in height) with weights attached to allow them to sink and rest on the pond bed. Traps were positioned in ponds as evenly as possible, with one central trap and four outer traps (e.g., Figure 1e). Traps were deployed for periods of up to 50 days before being retrieved, emptied and immediately redeployed. Collected sediment (including pond water pooled on the surface) from each trap was emptied into separate plastic bottles for transport back to the laboratory. Bottles were then emptied into larger plastic boxes and refrigerated for at least 48 h to allow suspended solids to settle out. The supernatant was then siphoned off into bottles and filtered following the same method described for SSC to account for the mass of any fine particles still in suspension. Macroinvertebrates found in trap samples were removed and identified to family level where possible. Sediment in the boxes was stirred thoroughly, and for each, three sub-samples of ~5 g were transferred into centrifuge tubes for particle size analysis. Grain size distributions and characteristics were determined using laser diffraction particle size analysis (Mastersizer 2000, Malvern Panalytical; Malvern, UK). Prior to analysis, samples were treated with a 5% sodium hexametaphosphate solution to disperse particles and agitated for 5 min in an ultrasonic bath. To determine sediment mass, the remaining sediment was distributed into pre-weighed aluminium trays (~100 g sediment per tray) and oven-dried at 105 °C for at least 48 h before being cooled and reweighed. To determine volatile solids (organic matter) by LOI, one tray per trap was then ignited at 500 °C for 2 h before being cooled and reweighed. One tray per batch was reheated and reweighed to check that the sample mass remained stable. P content was determined by grinding the ignited sample into a fine powder, of which triplicate subsamples of 3 ± 0.1 mg were taken, mixed with 60 mL ultrapure water and then analysed using the same TP methodology used for water samples. Length and width transects of pond sediment depths were surveyed in January and July 2020 following a standard method [46], and spatially interpolated in a GIS (ArcMap, Esri; Redlands, CA, USA) using the natural neighbour interpolation method to estimate stored sediment volumes. Measuring along transects aimed to minimise sediment disturbance and damage to habitat but meant that measurements were not evenly distributed across the pond area. Natural neighbour interpolation was, therefore, chosen over other methods because of its ability to perform well with an uneven sampling density and irregular distribution of data points [47].

2.4. Data and Statistical Analyses

Statistical procedures were carried out in RStudio v1.1.453 (RStudio Team, Boston, MA, USA, 2016) using the programming language R [48]. Inlet and outlet water quality determinand concentration data were assessed for normality with the Shapiro-Wilk test, after which any non-normal variables were normalised using cube-root transformations. Paired samples t-tests were carried out on inlet-outlet samples for determinands to compare their means over the sampling period [49]. Similarly, sediment particle size distribution variables were tested for normality and equal variances to ensure robustness before performing one-way ANOVAs and post hoc Tukey’s tests on the data [50]. Baseflow removal efficiencies of determinands from ponds were calculated using the following equation:

\[
\text{Removal Efficiency (\%)} = \frac{\text{Inflow Concentration} - \text{Outflow Concentration}}{\text{Inflow Concentration}} \times 100
\]  

(1)

Similarly, removal efficiencies for storm events were determined using total loads at monitoring locations calculated using the estimated event discharge. Antecedent Precipitation Index (API) was calculated for storm events with daily rainfall records since data collection began using the following equation:

\[
API_d = k \times API_{d-1} + P_d
\]

(2)
where API\_d is the API for day \(d\); \(k\) is a decay factor and \(P\_d\) is rainfall for day \(d\). A fixed value of 0.95 was used for \(k\) following the method described by Hill et al. (2015) [51]. Simple linear regressions were carried out to test the effect and strength of water temperature on determinand removal efficiencies where previous research had suggested the relationships exist [52,53].

### 3. Results

#### 3.1. Near Baseflow Water Quality

Outside of rainfall events, 19 sets of samples were taken between March 2019 and March 2020. Significant differences between inlet and outlet concentrations were found for dissolved nitrate, SRP; SSC and VSC, which all showed a decrease in mean concentration at the outlet (paired samples \(t\)-test, \(p < 0.01\), \(n = 19\)) (Figure 3).

The other determinands generally showed minimal variance between the inlet and outlet; however, in some cases TP/PP concentrations increased by over 100% at the outlet. Nitrite (NO\(_2^-\)) was excluded from the statistical tests due to a majority (67%) of both inlet and outlet samples measuring 0 mg NO\(_2^-\) L\(^{-1}\).

Removal efficiencies exhibited considerable variability between determinands during baseflows, ranging from extreme negative values (net export from the pond system) for PP, to more consistently positive values (net retention) for SSC and VSC (Table 1). Overall, the majority of mean removal efficiencies for the sampling period were positive, with the exceptions being PP, TP, and NH\(_4^+\).

<table>
<thead>
<tr>
<th>Determinand</th>
<th>Mean Inflow Concentration (mg L(^{-1}))</th>
<th>Mean Outflow Concentration (mg L(^{-1}))</th>
<th>Mean Removal Efficiency (%)</th>
<th>Minimum Removal Efficiency (%)</th>
<th>Maximum Removal Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRP</td>
<td>0.008 ± 0.006</td>
<td>0.005 ± 0.004</td>
<td>29 ± 37</td>
<td>-100</td>
<td>74</td>
</tr>
<tr>
<td>TDP</td>
<td>0.041 ± 0.023</td>
<td>0.038 ± 0.02</td>
<td>3 ± 43</td>
<td>-117</td>
<td>68</td>
</tr>
<tr>
<td>PP</td>
<td>0.04 ± 0.04</td>
<td>0.052 ± 0.059</td>
<td>-237 ± 579</td>
<td>-2100</td>
<td>95</td>
</tr>
<tr>
<td>TP</td>
<td>0.081 ± 0.048</td>
<td>0.089 ± 0.069</td>
<td>-34 ± 125</td>
<td>-314</td>
<td>77</td>
</tr>
<tr>
<td>NH(_4^+)</td>
<td>0.023 ± 0.026</td>
<td>0.024 ± 0.025</td>
<td>-61 ± 118</td>
<td>-400</td>
<td>73</td>
</tr>
<tr>
<td>NO(_3^-)</td>
<td>36.56 ± 3.585</td>
<td>34.903 ± 4.4</td>
<td>5 ± 6</td>
<td>-2</td>
<td>23</td>
</tr>
<tr>
<td>F(^-)</td>
<td>0.068 ± 0.023</td>
<td>0.067 ± 0.024</td>
<td>0 ± 18</td>
<td>-23</td>
<td>35</td>
</tr>
<tr>
<td>Cl(^-)</td>
<td>16.913 ± 2.382</td>
<td>16.835 ± 2.045</td>
<td>0 ± 7</td>
<td>-23</td>
<td>14</td>
</tr>
<tr>
<td>SO(_4^{2-})</td>
<td>17.006 ± 2.652</td>
<td>16.904 ± 2.488</td>
<td>0 ± 9</td>
<td>-29</td>
<td>18</td>
</tr>
<tr>
<td>SSC</td>
<td>21.2 ± 4.153</td>
<td>13.464 ± 6.943</td>
<td>32 ± 24</td>
<td>-17</td>
<td>70</td>
</tr>
<tr>
<td>VSC</td>
<td>7.09 ± 4.153</td>
<td>3.901 ± 1.469</td>
<td>40 ± 15</td>
<td>15</td>
<td>66</td>
</tr>
</tbody>
</table>

#### 3.2. Storm Event Water Quality

Four storm events were captured between March 2019 and February 2020 (Table 2); however, it was not always possible to trigger all four automatic samplers for every storm. The event captured in February was during Storm Dennis and had the highest rainfall; total monthly rainfall in February was 170% above average for the area. Estimated peak discharge was highest during the November event, with a return period of 5.5 years [54]. API was highest prior to the October 14th event following a rapid wetting of the catchment at the end of September.
Figure 3. Boxplots showing paired on-line pond inlet and outlet concentrations for (a) dissolved nitrate; (b) dissolved ammonium; (c) SRP; (d) TDP; (e) TP; (f) PP; (g) dissolved fluoride; (h) dissolved chloride; (i) dissolved sulfate; (j) SSC and (k) VSC. Median values are represented by bold lines. Significance levels for results of paired samples t-tests are indicated with: *** \((p < 0.001)\), ** \((p < 0.01)\), ns \((p > 0.05)\).

The March 2019 event was the smallest in magnitude, with the least rainfall and lowest API, but still resulted in a peak SSC of \(>200\) mg L\(^{-1}\) at the inlet to the Upstream Pond, with the peak then being reduced by \(~50\%\) downstream at the outlet of the Downstream Pond (Figure 4d). Streamflow responded rapidly to rainfall with a lag time of less than two hours (Figure 4a–b). The response of suspended sediment was partially staggered, with
lag times increasing downstream at each monitoring point except for water leaving the Downstream Pond, which peaked simultaneously with water leaving the Central Pond. SSC at the Downstream Pond outlet had a less steep gradient on the falling limb compared to the other monitoring locations. Volatile solids made up <20% of the total solids during peaks, but as high as 78% on the receding limb (Figure 4c).

Table 2. Mean (±SD) SSC (mg L\(^{-1}\)) for each pond monitoring site during four storm events, estimated discharge (L s\(^{-1}\)) prior to the event and at its peak, and the sampling duration (hours). Rainfall (mm) is the total event precipitation and Antecedent Precipitation Index (API) (mm) is given for the day prior to each event.

<table>
<thead>
<tr>
<th>Storm Event</th>
<th>Mean SSC (mg L(^{-1}))</th>
<th>Estimated Discharge (L s(^{-1}))</th>
<th>Rainfall (mm)</th>
<th>API (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upstream Pond Inlet</td>
<td>Upstream Pond Outlet</td>
<td>Central Pond Outlet</td>
<td>Downstream Pond Outlet</td>
</tr>
<tr>
<td>12/13 March 2019</td>
<td>45 ± 47</td>
<td>30 ± 33</td>
<td>29 ± 27</td>
<td>35 ± 30</td>
</tr>
<tr>
<td>14 October 2019</td>
<td>258 ± 365</td>
<td>161 ± 152</td>
<td>143 ± 94</td>
<td>126 ± 55</td>
</tr>
<tr>
<td>14 November 2019</td>
<td>92 ± 67</td>
<td>27 ± 11</td>
<td>24 ± 7</td>
<td>-</td>
</tr>
<tr>
<td>15/16 February 2020</td>
<td>-</td>
<td>87 ± 63</td>
<td>98 ± 79</td>
<td>-</td>
</tr>
</tbody>
</table>

The response of TP and PP closely reflected that of SSC and VSC; however, TDP did not exhibit a rising limb and remained relatively constant at the inlet and outlet of the Upstream Pond (Figure 4e–g). TDP showed a somewhat different pattern at the outlet of the Central Pond with the concentration abruptly dropping below 10 µg P L\(^{-1}\) after 19:00 p.m. At the Downstream Pond outlet, TDP remained under 20 µg P L\(^{-1}\), which was lower than both the inlet and outlet of the Upstream Pond which almost always stayed above 20 µg P L\(^{-1}\). On the rising and receding limbs of the event, PP accounted for the majority (57–91%) of transported P, after which TDP at the inlet and outlet of the Upstream Pond exceeded the particulate fraction. Automatic sampler SRP data are not presented as samples could not be analysed within 48 h of sampling and showed a 60–100% decrease when compared to grab samples analysed within 48 h. Grab samples showed that SRP made up to 42% of the TDP at 12:00 p.m.

Both dissolved ammonium and nitrate concentrations increased during the March event, with nitrate having more defined peaks and ammonium having a more variable response (Figure 4h,j). Dissolved nitrite displayed rising limbs at the pond outlets but remained comparatively low at the Upstream Pond inlet (Figure 4i). Throughout the event, the majority of dissolved N transported was made up by nitrate. Concentrations of nitrate after the peak remained consistently higher leaving the Upstream Pond than those at the inlet. Dissolved fluoride showed a rising limb during the storm event after which the concentration decreased gradually and returned to a similar level as at the start of the event (Figure 4m). Dissolved chloride and sulfate concentrations exhibited almost identical patterns, with both solutes showing a small dilution between 14:00 p.m. and 15:00 p.m. coinciding with the peak water level (Figure 4k,l). There was minimal variation in chloride and sulfate concentration between sampling sites with the exception of two sudden peaks at the outlet of the Upstream Pond.
Figure 4. Timeseries during a storm event on 12/13 March 2019 showing: (a) hourly rainfall (mm); (b) stage (m) in the Central Pond; and concentrations of water quality determinands: (c) VSC and (d) SSC (mg L$^{-1}$); (e) TP, (f) PP, and (g) TDP ($\mu$g P L$^{-1}$); (h) ammonium (mg NH$_4^+$ L$^{-1}$); (i) Nitrite (mg NO$_2^-$ L$^{-1}$); (j) Nitrate (mg NO$_3^-$ L$^{-1}$); (k) chloride (mg Cl$^-$ L$^{-1}$); (l) Sulfate (mg SO$_4^{2-}$ L$^{-1}$) and (m) Fluoride (mg F$^-$ L$^{-1}$) at each pond inlet/outlet sampling site.
During the sampled storm events, total suspended sediment loads entering the pond system varied from 55 to 220 kg, and between 0.08 and 0.44 kg for TP, reflecting both the event magnitude and duration (Table 3). Load removal efficiencies varied greatly between ponds and events; however, the Upstream Pond was consistently the most efficient in all events sampled for both suspended sediment and TP. Generally, load removal efficiencies were higher for suspended sediment than TP, with negative removal efficiencies occurring more frequently for TP. During the March event, the Downstream Pond showed the lowest (negative) removal efficiency for suspended sediment out of all sampled events as a result of elevated concentrations at its outlet during the falling limb. The highest removal efficiency was observed in the November event for the Upstream Pond which indicated a net retention of 50 kg suspended sediment during <6 h. Sediment load removal efficiency of the Upstream Pond in the October event was 40% lower, but still retained 66 kg also across a <6-h period. Overall sediment retention in the March event was comparatively much smaller at only 9 kg over a longer 23-h period. In the context of the wider catchment area, the net sediment load retained by the ponds in the March event was equivalent to 0.85% of the flux leaving the Downstream Catchment during the same time period. This proportion was almost 2.5 times greater during the October event, with 2.1% of the Downstream Catchment flux retained by the ponds.

3.3. Pond Sediment Quality

From the manual surveying of sediment depths approximately two years after their construction, it was estimated that 13.89 m$^3$ of matter had accumulated in the Upstream Pond and 7.36 m$^3$ in the Central Pond. This meant that the Upstream Pond had filled ~20% of its total capacity, and the Central ~8%. At the time of surveying in January, depths in the Downstream Pond were unable to be measured due to the water level being too high. The Downstream Pond was able to be surveyed in July at the earliest (due to the Covid-19 pandemic), and had accumulated 9.89 m$^3$ of matter, equating to ~10% of its total capacity.

Sediment traps were first deployed in March 2019, after which traps were deployed continuously from August 2019 with sediment collection taking place on six occasions until March 2020 to capture run-off during the wet season. Throughout this seven month period, rates of accumulation were variable, but the Upstream Pond had the highest overall accumulation, and the Downstream Pond had the lowest (Table 4). Sediment accumulation rates varied considerably between the trap placements within ponds as shown by the large standard deviations. Despite only a short deployment period, the accumulations were surprisingly high during August, with ponds accumulating disproportionately more sediment (0.048 t ha$^{-1}$) than the yield leaving the Downstream Catchment (0.001 t ha$^{-1}$). Over the whole period, the ponds accumulated 6.1% of the downstream catchment silt + clay flux, and 7.6% of all suspended sediment. P accumulation in ponds generally showed the same pattern as sediment, and on average made up ~0.1% of the total accumulated mass (Table 5). Total accumulated P in ponds only made up 3.2% of the Downstream Catchment P flux. LOI showed that deposited sediments were largely made up of inorganic matter (IOM), accounting for >75% of the accumulated sediment mass throughout the sampling period. The organic matter (OM) content ranged from 10–23% and consistently decreased downstream along the pond sequence in each deployment period. OM content was highest between August and October. OM content of pond sediment was significantly enriched compared to the soil in the arable fields of the contributing area, which had an OM content of 5–7%, typical of arable fields in this area.
Table 3. Load (kg) of suspended sediment, TP at each monitoring location and removal efficiency (%) of each pond for the sampled storm events. Ups. = Upstream Pond; Cent. = Central Pond; Down. = Downstream Pond; Catch. = gauged catchment. (NB Loads for TP in the March 2019 event are calculated for the first half of the event period).

<table>
<thead>
<tr>
<th>Storm Event</th>
<th>Suspended Sediment</th>
<th>TP</th>
<th>Suspended Sediment</th>
<th>TP</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 2019</td>
<td>54.73</td>
<td>38.9</td>
<td>37.81</td>
<td>46.21</td>
</tr>
<tr>
<td>October 2019</td>
<td>219.96</td>
<td>154.05</td>
<td>141.80</td>
<td>126.94</td>
</tr>
<tr>
<td>November 2019</td>
<td>70.55</td>
<td>20.53</td>
<td>17.81</td>
<td>-</td>
</tr>
<tr>
<td>February 2020</td>
<td>-</td>
<td>213.5</td>
<td>250.79</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4. Accumulated sediment (±SD) (t) in each pond, all three ponds, and only the silt + clay (<63 µm) for sediment trap monitoring periods. Accumulated sediment yield (t ha⁻¹) for all ponds from the contributing area (30 ha), the flux of sediment and silt + clay (t) and the exported yield (t ha⁻¹) from the downstream catchment area (340 ha) are given for the same periods.

<table>
<thead>
<tr>
<th>Monitoring Period</th>
<th>Days</th>
<th>Rainfall (mm)</th>
<th>Accumulated Sediment (t)</th>
<th>All Ponds (silt+clay)</th>
<th>All Ponds Sediment Yield (t ha⁻¹)</th>
<th>Catchment Sediment Flux (t)</th>
<th>Catchment silt-clay Flux (t)</th>
<th>Catchment Sediment Yield (t ha⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 August 2019–30</td>
<td>22</td>
<td>62</td>
<td>0.56 ± 0.27</td>
<td>0.54 ± 0.35</td>
<td>0.33 ± 0.35</td>
<td>1.43 ± 0.56</td>
<td>1.01</td>
<td>0.048</td>
</tr>
<tr>
<td>August 2019</td>
<td></td>
<td></td>
<td>0.54 ± 0.35</td>
<td>0.33 ± 0.35</td>
<td>1.43 ± 0.56</td>
<td>1.01</td>
<td>0.048</td>
<td>0.34</td>
</tr>
<tr>
<td>30 August 2019–3</td>
<td>34</td>
<td>128</td>
<td>0.63 ± 0.55</td>
<td>0.17 ± 0.05</td>
<td>0.25 ± 0.04</td>
<td>1.06 ± 0.55</td>
<td>0.71</td>
<td>0.035</td>
</tr>
<tr>
<td>October 2019</td>
<td></td>
<td></td>
<td>0.63 ± 0.55</td>
<td>0.17 ± 0.05</td>
<td>0.25 ± 0.04</td>
<td>1.06 ± 0.55</td>
<td>0.71</td>
<td>0.035</td>
</tr>
<tr>
<td>3 October 2019–30</td>
<td>27</td>
<td>132</td>
<td>0.69 ± 0.27</td>
<td>0.32 ± 0.11</td>
<td>-</td>
<td>1.01 ± 0.29</td>
<td>0.65</td>
<td>0.034</td>
</tr>
<tr>
<td>October 2019</td>
<td></td>
<td></td>
<td>0.69 ± 0.27</td>
<td>0.32 ± 0.11</td>
<td>-</td>
<td>1.01 ± 0.29</td>
<td>0.65</td>
<td>0.034</td>
</tr>
<tr>
<td>30 October 2019–4</td>
<td>35</td>
<td>140</td>
<td>0.63 ± 0.37</td>
<td>0.39 ± 0.2</td>
<td>-</td>
<td>1.02 ± 0.42</td>
<td>0.67</td>
<td>0.034</td>
</tr>
<tr>
<td>December 2019</td>
<td></td>
<td></td>
<td>0.63 ± 0.37</td>
<td>0.39 ± 0.2</td>
<td>-</td>
<td>1.02 ± 0.42</td>
<td>0.67</td>
<td>0.034</td>
</tr>
<tr>
<td>4 December 2019–22</td>
<td>49</td>
<td>167</td>
<td>0.67 ± 0.27</td>
<td>0.82 ± 0.28</td>
<td>0.67 ± 0.23</td>
<td>2.16 ± 0.45</td>
<td>1.57</td>
<td>0.072</td>
</tr>
<tr>
<td>January 2020</td>
<td></td>
<td></td>
<td>0.67 ± 0.27</td>
<td>0.82 ± 0.28</td>
<td>0.67 ± 0.23</td>
<td>2.16 ± 0.45</td>
<td>1.57</td>
<td>0.072</td>
</tr>
<tr>
<td>22 January 2020–12</td>
<td>50</td>
<td>177</td>
<td>0.98 ± 0.35</td>
<td>1.05 ± 0.33</td>
<td>0.49 ± 0.31</td>
<td>2.52 ± 0.57</td>
<td>1.77</td>
<td>0.084</td>
</tr>
<tr>
<td>March 2020</td>
<td></td>
<td></td>
<td>0.98 ± 0.35</td>
<td>1.05 ± 0.33</td>
<td>0.49 ± 0.31</td>
<td>2.52 ± 0.57</td>
<td>1.77</td>
<td>0.084</td>
</tr>
<tr>
<td>Total</td>
<td>217</td>
<td>871</td>
<td>4.15 ± 0.89</td>
<td>3.29 ± 0.6</td>
<td>1.74 ± 0.52</td>
<td>9.18 ± 1.19</td>
<td>6.38</td>
<td>0.306</td>
</tr>
</tbody>
</table>

| | |
| | |
Table 5. Accumulated phosphorus (±SD) (kg) in each pond and all three ponds for sediment trap monitoring periods. Accumulated P yield (kg ha$^{-1}$) for all ponds from the contributing area (30 ha), the flux of P (kg) and the exported P yield (kg ha$^{-1}$) from the downstream catchment area (340 ha) are given for the same periods.

<table>
<thead>
<tr>
<th>Monitoring Period</th>
<th>Days</th>
<th>Rainfall (mm)</th>
<th>Accumulated P (kg)</th>
<th>All Ponds P Yield (kg ha$^{-1}$)</th>
<th>Catchment P Flux (kg)</th>
<th>Catchment P Yield (kg ha$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 August 2019–30 August 2019</td>
<td>22</td>
<td>62</td>
<td>0.58 ± 0.27</td>
<td>0.51 ± 0.34</td>
<td>0.29 ± 0.28</td>
<td>1.38 ± 0.52</td>
</tr>
<tr>
<td>30 August 2019–3 October 2019</td>
<td>34</td>
<td>128</td>
<td>0.69 ± 0.55</td>
<td>0.22 ± 0.08</td>
<td>0.27 ± 0.04</td>
<td>1.18 ± 0.56</td>
</tr>
<tr>
<td>3 October 2019–30 October 2019</td>
<td>27</td>
<td>132</td>
<td>0.65 ± 0.18</td>
<td>0.36 ± 0.12</td>
<td>-</td>
<td>1.01 ± 0.22</td>
</tr>
<tr>
<td>30 October 2019–4 December 2019</td>
<td>35</td>
<td>140</td>
<td>0.56 ± 0.29</td>
<td>0.4 ± 0.19</td>
<td>-</td>
<td>0.96 ± 0.35</td>
</tr>
<tr>
<td>4 December 2019–22 January 2020</td>
<td>49</td>
<td>167</td>
<td>0.6 ± 0.22</td>
<td>0.81 ± 0.27</td>
<td>0.69 ± 0.25</td>
<td>2.1 ± 0.43</td>
</tr>
<tr>
<td>22 January 2020–12 March 2020</td>
<td>50</td>
<td>177</td>
<td>0.91 ± 0.22</td>
<td>0.94 ± 0.48</td>
<td>0.42 ± 0.27</td>
<td>2.27 ± 0.59</td>
</tr>
<tr>
<td>Total</td>
<td>217</td>
<td>871</td>
<td>3.99 ± 0.77</td>
<td>3.24 ± 0.69</td>
<td>1.68 ± 0.46</td>
<td>8.91 ± 1.13</td>
</tr>
</tbody>
</table>
Total P content of sediment from traps varied from 695 to 1634 mg kg\(^{-1}\) (Figure 5). The highest median P content in each pond occurred during September and then showed a downward trend in the following months. During most deployment periods, P content decreased along the pond sequence.

![Figure 5](image-url)

**Figure 5.** Boxplots showing the range of phosphorus content (mg kg\(^{-1}\)) of deposited sediment in each pond for each trap deployment period. Median values are represented by bold lines.

The P content of pond sediment was found to be positively correlated with OM content (\(p < 0.05\)) (Figure 6). This relationship was strongest in the Central Pond, with OM explaining 54% of variation in P content.

![Figure 6](image-url)

**Figure 6.** Linear regressions of organic matter content (%) and P content (mg kg\(^{-1}\)) of deposited sediment in each pond (Upstream Pond \(n = 34\); Central Pond \(n = 35\); Downstream Pond \(n = 20\)).
Deposited sediment in all three ponds was mainly comprised of the silt fraction, followed by sand and then clay, which only accounted for up to 6% of particles (Figure 7). Both clay and silt content showed an increasing trend downstream along the pond sequence, whilst sand showed a decrease. All pairwise comparisons show significant differences between group means with the exception of clay content between the Central and Downstream Ponds. Soil in the contributing area is known to have a clay content of 10–25%, silt content of 50–60%, and sand content of 15–30%, broadly mirroring the composition of the deposited pond sediments.

![Boxplots showing the range of sediment grain size content (%) of (a) sand; (b) silt and (c) clay in each pond from trap deployment during March 2019 (Upstream Pond n = 4; Central Pond n = 5; Downstream Pond n = 5). Median values are represented by bold lines and significance levels for Tukey’s tests by **** (p < 0.0001), *** (p < 0.001), ** (p < 0.01), * (p < 0.05), ns (p > 0.05).](image)

Median particle diameter (D_{50}) of deposited sediment in traps was shown to decrease downstream along the pond sequence, ranging from a maximum of 61.85 µm in the Upstream Pond to a minimum of 15.8 µm in the Downstream Pond (Figure 8a). Inversely, specific surface area (SSA) increased along the pond sequence, ranging from 0.42 m² g⁻¹ in the Upstream Pond to 0.89 m² g⁻¹ in the Downstream Pond (Figure 8b). Pairwise comparisons showed significant differences (p < 0.05) in both D_{50} and SSA between the Upstream and Central Pond, and the Upstream and Downstream Pond.

The P content of accumulated sediment within ponds showed much less variation compared to the P content of suspended sediment sampled during the March 2019 storm event (Figure 9). P content of suspended samples varied from 0 to ~2500 mg kg⁻¹. The suspended samples showed a general increasing trend in P content from upstream to downstream along the pond sequence; however, this enrichment effect appeared to level off between the Central Pond Outlet and Downstream Pond Outlet.
event (Figure 9). P content of suspended samples varied from 0 to ~2500 mg kg\(^{-1}\) compared to the P content of suspended sediment sampled during the March 2019 storm. This is considerably lower than the average nitrate removal efficiencies of between 30% and 87% reported by previous studies [21,52,53]. SRP removal efficiency of the online pond system had a wider range of up to 74%, though it is important to note that this increase during the warm seasons [23,55]. Other studies have found average SRP (or orthophosphate) removal efficiencies of between 12% and 87% [21,56,57]. The removal efficiency of nitrate showed increases during the warm seasons [23,55].

The on-line ponds were shown to be effective at removing dissolved nitrate and SRP, which are both bioavailable forms of N and P. The removal efficiency of nitrate showed seasonality, peaking during the summer at 23% but was below 10% the rest of the time, and only exhibited negative removal efficiencies on two occasions (in December and January). This is considerably lower than the average nitrate removal efficiencies of

4. Discussion

4.1. Near Baseflow Water Quality

The on-line ponds were shown to be effective at removing dissolved nitrate and SRP, which are both bioavailable forms of N and P. The removal efficiency of nitrate showed seasonality, peaking during the summer at 23% but was below 10% the rest of the time, and only exhibited negative removal efficiencies on two occasions (in December and January). This is considerably lower than the average nitrate removal efficiencies of
between 72% and 83% reported by previous studies [21,52,53]. SRP removal efficiency of the on-line pond system had a wider range of up to 74%, though it is important to note that this represents a reduction of only ~15 μg P L⁻¹ due to low concentrations of SRP in the baseflow. In contrast to nitrate, SRP removal efficiency did not show any apparent seasonality, though the lowest removal efficiency of −100% was also observed in winter. Evidence from a created wetland in Ohio, also draining an agricultural catchment, found no significant difference in removal between seasons; however, the influence of season on SRP removal has been shown to be important in certain wetlands, which exhibit removal increases during the warm seasons [23,55]. Other studies have found average SRP (or orthophosphate) removal efficiencies of between 12% and 87% [21,56,57]. The removal efficiency of bioavailable nutrient fractions is likely limited within the on-line ponds by several factors, a key one being the abundance and density of macrophytes and algae. During the first growing season (spring/summer 2018), limited establishment of vegetation was observed in the ponds. This could partly be due to extreme temperatures (2018 being the hottest summer on record in England), but also potentially a result of not enough time for natural colonisation to occur. In contrast, by the end of summer 2019 the ponds had been partly colonised, particularly the Central Pond with several stands of *Typha latifolia* and *Juncus* spp. In both years, all three ponds showed substantial algal growth, often forming thick mats of filamentous green algae that covered up to a quarter of pond surfaces. As data collection only began in 2019, the effect of the presence of vegetation on nutrient removal could not be analysed. However, it is thought that, over time, the ponds are likely to increase their nutrient removal capacity given further macrophyte succession. This is supported by a review of constructed wetlands that found that SRP and TP removal was higher in older (>18 months) wetlands [30]. SRP removal is also influenced by the underlying pond bed sediment and its P sorption/desorption capacity, typically quantified as the Equilibrium Phosphorus Concentration (EPC₀) [58]. However, in constructed wetlands with considerable algal growth, EPC₀ has been shown to play a less important role in P removal compared to algal uptake [59]. Our study did not undertake EPC₀ measurements for consideration of SRP removal; however, bed sediment P enrichment in ponds was measured using the sediment traps (Section 4.3). SRP retention in the ponds may also be aided by the persistently high nitrate concentrations, which can buffer the reductive dissolution of Fe and thereby limit any redox-mediated SRP release [60–62].

An important biological nutrient removal mechanism for N in wetlands is bacterial metabolism, most commonly through nitrification and denitrification pathways [63]. In our study, dissolved ammonium concentrations showed no significant difference between the inlet and outlet, suggesting that nitrification is an unlikely cause of N removal in these ponds. However, denitrification is more likely to be occurring in the ponds to reduce nitrate to nitric oxide, nitrous oxide, and nitrogen gas [64]. Previous studies showed that denitrification rates are increased under anoxic or low dissolved oxygen conditions, warmer temperatures and an optimum pH of between 6 and 8 [65]. A positive correlation between mean daily water temperature in the Central Pond and nitrate removal efficiency (linear regression, R² = 0.32, p = 0.06) supports these findings. Temperature was only able to explain almost a third of the variation in removal efficiency, but this is justifiably low due to the other influential factors mentioned above not being considered. Nitrate removal in a constructed wetland in North Carolina showed a similar temperature dependence, with removal efficiencies of >90% during the growing season [53]. Although our monitoring showed reduced levels of nitrate removal during the winter, the net losses from the ponds during this period were minimal (<2% concentration increase in the outflow) and, therefore, not significantly affecting water quality. Further monitoring of the ponds (particularly dissolved oxygen measurements) may provide further data to help explain nutrient removal efficiencies and processes.

PP was not significantly reduced at the outlet, but both SSC and VSC were and had removal efficiencies of up to 70% and 66% respectively. It was hypothesised that the
majority of the inflowing P load would be sediment-bound and settle out in the ponds but results show that PP in the outflow remained just as high. A potential explanation could be that under lower flows, there was an increased export of planktonic algae from their proliferation in the ponds, as well as clay particles that remained suspended during low flow velocities. A study on stormwater control structures found evidence of PP release during low flows which were attributed to SRP release from anaerobic sediment, which was then adsorbed onto clay particles or assimilated by algae [15]. The observed reduction in SSC and VSC was expected due to the rapid reduction in flow velocity within ponds, which likely resulted in the deposition of larger particles with higher settling velocities nearest the inflow.

4.2. Storm Event Water Quality

The ponds were most effective at reducing suspended solids (SSC and VSC) downstream, and to a lesser extent P. The ponds may have a lower removal efficiency for TP/PP compared to SSC/VSC due to a high proportion of the P being bound to clay particles, which are more likely to remain in suspension compared to the coarser-grained particles. Studies show that the particulate fraction of P is often adsorbed onto the surface of particles such as metal oxides (e.g., iron oxides), or on clay particles [66]. Although water samples were unable to be analysed for particle size, the downstream increase in P content of suspended sediment in storm samples suggests that heavier particles settled out in the Upstream and Central Ponds, thereby reducing SSC, but having a smaller effect on PP concentrations, which are more heavily influenced by finer particles. Generally, SSC was able to explain most of the variation in suspended PP concentration, particularly under storm conditions (Figure 10).

![Graph showing linear regression of SSC and PP](image)

**Figure 10.** Linear regression of SSC (mg L$^{-1}$) and PP (mg P L$^{-1}$) from Pond Inlet/Outlet sampling sites during a storm event on the 12th and 13th March 2019, and near baseflow from the Inlet and Outlet ($n = 96$). The grey band represents 95% confidence intervals.

It is thought that under near baseflow conditions, PP at the outlet tended to be high despite the low SSC because the finest particles, with much lower masses remaining in suspension even during very low flow velocities. The ~1:1 width-to-length ratios of the ponds are likely to have limited their potential to remove PP and clay sized particles, with previous work suggesting that width-to-length ratios play an important role in controlling hydraulic efficiency and pollutant removal efficiency [67]. Longer ponds, with width-to-length ratios of greater than 1:4, have increased flow pathways and may, therefore,
have improved ability to settle out and retain finer sediment particles and, consequently, PP [68,69].

Particulate forms of P are also associated with organic P compounds (e.g., organophosphates used as pesticides), altogether making this fraction chemically and physically complex with highly variable stabilities, bondings and exchangeabilities [70]. A study modelling agricultural best management practices found that ponds were more effective at removing organophosphates, such as chlorpyrifos, that are more readily attached to sediment particles [71]. Sediment-bound P can be released into the water column through resuspension and desorption, which is more likely to occur during high magnitude storm events. This phenomenon is also likely to explain the increase in suspended sediment and TP loads downstream of the Central Pond during the large February 2020 event. Furthermore, during multiple events, both the Upstream and Central Pond were seen to be partially overtopping their banks when outflows were not able to drain fast enough to accommodate the inflowing discharge.

Barber and Quinn (2012) found that during a storm event an on-line run-off attenuation feature was not able to reduce SSC, TP or nitrate by significant levels at its outlet [13]. In terms of sediment and P they attributed these findings to the resuspension of previously deposited material, highlighting this as a key drawback of interventions of this style. Our study also found that nitrate was not retained during the storm event in March 2019, but instead appears to have been flushed out of the Upstream Pond at a higher concentration.

Suspended material entered the pond system at notably higher concentrations during the October and February events, where it was observed that a substantial overland flow pathway located just upstream of the Upstream Pond was active. During the October event, run-off from this pathway (shortly after the storm’s peak) contributed an SSC of 219 mg L\(^{-1}\) and a TP concentration of 0.75 mg L\(^{-1}\), whilst SSC at the inflow was approximately four times higher, with a TP concentration of 1.1 mg L\(^{-1}\). It is likely that a combination of antecedent conditions and intense rainfall brought about the activation of this critical source area to significantly increase stream sediment delivery from the hillslope.

4.3. Pond Sediment Quality

Sediment particle size is a key parameter in determining the transport and fate of pollutants in streams [72]. The chief concerns for water quality are the smallest sediment particles (clay and fine silt) that are capable of transporting large quantities of bound P when entrained and remain in suspension for longest [69]. High concentrations of fine sediment typically result in turbid water and have been shown to have adverse ecological consequences (both in suspension and when deposited) for primary productivity, aquatic food webs, benthic macroinvertebrate communities and salmonid spawning habitats [73–76]. Results from the sediment traps show that it was mostly the silt fraction (2–63 µm) being deposited in all three ponds; however, there was a shift towards a smaller median particle size along the sequence. This is largely a consequence of the deposition and filtering out of sand particles within upstream ponds which significantly decreases their proportion within sediment downstream. Visual observations showed that there was considerable build-up of coarse matter at pond inlets. Similar results were found by Cooper et al. (2019), who showed a decrease in the mean particle size of deposited sediment along the length of a constructed wetland [14]. Comparable results were found by Ockenden et al. (2014) who also observed that in paired field wetland ponds, the median particle size was typically larger in the first pond, which reflects the results of our study [77]. They also found that sediment nutrient concentrations were generally higher in the second pond of the pair. In our study, the observed increase in SSA along the pond sequence would suggest that the Downstream Pond may have a higher capacity to trap P, given the importance of particulate surface area for adsorption of sediment-associated contaminants [78,79]. However, our data do not support this idea and, in fact, show the opposite relationship, suggesting that particle size characteristics are not the dominant influence on sediment-bound P within this system. A potential explanation for this may be that organic matter within ponds
has a greater contribution to sediment P enrichment. Sediment P content appeared to show a seasonal pattern that peaked during September. The increase in P content with organic matter suggests that a significant proportion of P within the pond sediment was derived from plant material, with the strongest correlation seen in the Central Pond. A similar relationship was observed in riverbed sediment from the River Blackwater, where organic matter and iron content explained 59% of variation in P [80]. During the autumn, there was a noticeable increase in leaf litter found within sediment traps, particularly in the Upstream Pond which is immediately downstream of a ~350 m length of stream with dense riparian tree cover. Decomposition of the macrophytes within ponds is likely to have been a key source of autochthonous organic matter and P, particularly in the Central Pond where macrophyte cover was greatest. It is likely that the ponds may have different rates of internal P cycling processes such as the release of dissolved P from sediment back into the water column as a result of organic matter breakdown and mineralization [81]. This dissolved P may also accumulate within interstitial pore water in the pond sediment, allowing it to be assimilated by rooted macrophytes, e.g., Typha, or potentially released into the water column if significant disturbance and remobilisation occurs as a result of a storm event [82,83].

The ratios of PP to suspended sediment at both the inlet and outlet of the Upstream Pond during the March storm event were over double those of deposited sediment within the traps. However, the ratio at the inflow during near baseflow conditions was 20% lower. These relationships between sediment and PP content suggest that there may be release of P from sediment within the pond during events, similar to the findings of Barber and Quinn (2012) [13], whereas the sediment is more likely to become enriched in P under average flow conditions.

4.4. Pond Capacity

The ponds showed significant accumulation during their first two years of being in operation, with an estimated annual reduction in capacity of 10% for the Upstream Pond, and 5% for both the Central and Downstream Pond. The storm event sampling data suggest that loads of up to 66 kg sediment can be retained during stormflows in the absence of flushing; however, up to 37 kg could be lost when flushing occurs. Further establishment of pond vegetation is likely to reduce the risk of flushing, with previous evidence demonstrating a decrease in sediment resuspension with increasing macrophyte cover [84]. Continued monitoring over successive years is needed to investigate this effect.

Without further intervention to maintain storage capacity, it is also possible that the ponds may undergo periods of net accumulation followed by net export if event magnitude is sufficiently large. In light of this, ponds may be even more prone to flushing in the future with climate change predicted to intensify extreme precipitation events and flood risk [85]. In order to overcome and limit the issue of remobilisation and flushing of accumulated matter downstream during high magnitude events, regular maintenance will be required. Ponds can be dredged most efficiently during periods of low flow in summer when water levels are minimal. The first pond within a sequence will need dredging more frequently (at least every two years) than the following ponds. There should also be consideration of the impacts of dredging on pond ecology where maintenance activities may damage habitat and remove vegetation. Keeping sections of vegetated sediment intact will aid recolonisation and reduce the risk of resuspension following maintenance. In practice, maintenance frequency will likely be a trade-off between its cost and the effectiveness of the ponds as mitigation measures.

4.5. Ecology

During the ~2.5 years since creation, the stream reach, ponds and marginal areas were colonised from bare soil into wetland habitat, with a plant species richness of 31 as of August 2020. A range of benthic macroinvertebrate taxa were also recorded throughout the monitoring period from a total of 22 different families. The presence of filamentous
green algae was observed within all ponds throughout the monitoring period, likely due to a lack of shading and persistently high nitrate concentrations.

In terms of the potential ecological impact of P being exported from the ponds, it can be said the risk for eutrophication downstream is low due to the flushing phenomenon being observed during the winter period when flows are typically high enough for sufficient dilution of P. Intense convective storms during summer are likely to pose a greater risk for eutrophication, though their occurrence is less frequent. In terms of fine sediment flushing, there are potential risks of contributing to benthic smothering, since during winter many benthic spawning biota have eggs in the riverbed. Further research into the ecological impacts of pond features on downstream communities would be beneficial for a more holistic evaluation of overall costs and benefits, particularly if monitored over a longer timescale.

4.6. Implications for Catchment Management

Trapping effectiveness was highly variable across the monitoring period for different water quality determinands and hydrological conditions. The surprisingly high accumulation of sediment in the ponds compared to the downstream catchment flux during August may be a result of several convective storms occurring during this period. These events may have been considerably localised, thereby mobilising sediment upstream of the ponds, but only having minimal impact on sediment transport in the rest of the catchment. In the context of the wider 340 ha catchment, the total accumulations in the ponds made up significant proportions within the overall budget (7.6% of suspended sediment; 6.1% of silt and clay; 3.2% of P) given that the ponds only drain 8.8% of the Downstream Catchment. It is important to note that the proportion of the flux trapped by ponds is likely to represent an upper estimate, because not all of the sediment would have necessarily been transported to the Downstream Catchment Outlet, particularly the larger particles. The estimated flux of clay and silt is, therefore, a more realistic representation of the suspended load exported from the catchment. Our findings highlight how the ponds show most potential for reducing downstream sediment loads, but are less efficient for mitigating diffuse agricultural P pollution. Despite only covering a small area (<0.02%) of the wider catchment, the ponds trapped a disproportionately large percentage of the fine sediment flux leaving the catchment. This highlights the importance of locating ponds where they will intercept high yielding run-off pathways within the catchment, and also makes them a particularly beneficial mitigation intervention where space is limited, and it is not economically viable for farms to lose large areas of agriculturally valuable land. Currently in the UK (and under the WFD) there are no regulatory limits on fluvial suspended sediment concentrations or yields. Without robust and specific sediment targets, the estimated pond sediment accumulations are difficult to assess in terms of ecological and regulatory significance; nevertheless, such interventions show useful potential as management tools in the delivery of on-farm pollution mitigation.

This paper provides further evidence on how the trapping efficiency of in-stream pond features is often dependent on the magnitude and frequency of storm events they experience, with high discharge and sediment inputs leading to a rapid reduction in storage capacity and causing ponds to overflow. These issues could potentially be alleviated by altering pond designs to allow greater storage capacity or incorporating additional features (e.g., vegetated swales, woody debris dams) to capture and filter overflow. It is important to note that the young age of the ponds may also play a role in their limited ability to remove pollutants such as fine sediments. The expectation from catchment management efforts is often that observable benefits in pollutant reductions will be delivered shortly following implementation. The evidence presented here only shows a ‘snapshot’ of the ponds’ functioning and trapping efficiency in the short term, and it is very likely to change with continued geomorphological evolution of the stream channel and further colonisation and succession of vegetation. Continued monitoring would be beneficial for evaluating the ponds’ performance over a time period that allows for maintenance
and revegetation to take place. The capacity reduction of the ponds observed during this two-year period necessitates regular maintenance and poses the potential opportunity for disposal of deposited pond sediment back into the landscape. The sediment has value for farmers that can capitalise on its nutrient content by redistributing it on arable fields as a soil conditioner, though critical source areas should be avoided to minimise the risk of mobilisation following application. The accumulated pond sediment properties show good suitability for agricultural application, having high organic matter and silt content (silty loam texture) and thus good water holding capacity. Previous research demonstrates that dredged fluvial sediments can increase crop productivity if added to soil with poor agricultural characteristics, for example where soil organic matter has been depleted [86].

Even with the implementation of on-line pond features in agricultural headwaters, the delivery of other mitigation interventions and sustainable management practices are still required to enable the best chance of achieving ecologically significant improvements to water and habitat quality in downstream catchments [87]. The value of the wider co-benefits from pond features is also important to consider in their overall evaluation and contribution to achieving catchment management objectives related to habitat and the aesthetic quality of the landscape. Monitoring of water quality and ecology can help assess benefits and risks postimplementation, thereby informing decisions on adaptive management for improving interventions.

5. Conclusions

This paper demonstrates how the effectiveness of on-line ponds for the mitigation of diffuse agricultural pollution on clay soils with a 2.5% slope can be highly variable due to the different retention capacities of sediment and nutrient fractions under different hydrological conditions. During baseflows, ponds reduced dissolved nitrate and SRP concentrations by averages of 29% and 5%, respectively. Despite their small size (<0.05 ha) and contributing area (30 ha), the on-line ponds were capable of accumulating significant pollutant masses over a seven-month period, equating to 7.6%, 6.1% and 3.2% of the wider catchment (340 ha) suspended sediment, silt and clay, and P fluxes, respectively. However, data suggest that net losses of sediment and P can occur during higher magnitude storm events, with this risk likely to increase as pond storage capacity reduces. The ponds are most advantageous for capturing silt and sand-sized material during smaller to medium events typically experienced during winter. This design of on-line pond with a ~1:1 width-to-length ratio is less effective at mitigating TP loading. We recommend that pond maintenance should be considered on a biennial basis, and removed sediment be reapplied to arable land as an organic-rich soil conditioner. In addition to the implications for water quality, these interventions provide benefits for habitat diversity and potential for flood attenuation in NFM schemes. On-line ponds are likely to be most effective when they are well-managed and used in combination with other mitigation measures, particularly helping to improve functioning during more extreme storm events. Further research into the longer-term evolution of the on-line pond system would help evaluate changes in its functioning over time with continued development of its geomorphology and vegetation.

Author Contributions: Conceptualization, J.R., G.O., P.R. and D.S.; methodology, J.R., G.O., P.R., D.S. and D.G.-T.; field work, J.R. and J.B.; data analysis, J.R.; investigation, J.R.; resources, G.O. and P.R.; writing—original draft preparation, J.R.; writing—review & editing, G.O., D.S., P.R., D.G.-T., J.B., J.O. and D.M.; visualization, J.R.; supervision, G.O., D.S., P.R. and D.G.-T.; funding Acquisition, G.O. and D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Environment Research Council, grant number NE/L002531/1. Additional research funding was also provided by Thames Water as part of J.R.’s PhD CASE studentship.

Institutional Review Board Statement: Not applicable.
Data Availability Statement: Data will be made publicly available through the Environmental Information Data Centre (EIDC) and subject to an embargo period ending in 2022 after data have been used in the PhD thesis of J.R.

Acknowledgments: We would like to thank the Astor family and farm manager Matt Childs for their permission and enthusiasm in undertaking this study on the Bruern Estate. Thank you to the Environment Agency and Wild Oxfordshire (Ann Berkeley) for the opportunity to monitor their pilot NFM scheme, and to Atkins, Thames Water (Helena Soteriou) and other members of the Evenlode Catchment Partnership for their support in helping to facilitate this research. We are also grateful to the Parish Council and community of Milton-under-Wychwood for their support and engagement with this research. Lab assistance and guidance on sample analysis from Mike Bowes, David Nicholls and Linda Armstrong is greatly appreciated.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References
15. Duan, S.; Newcomer-Johnson, T.; Mayer, P.; Kaushal, S. Phosphorus retention in stormwater control structures across streamflow in urban and suburban watersheds. Water 2016, 8, 390. [CrossRef]


60. Dupas, R.; Musolff, A.; Jawitz, J.W.; Suresh, P.; Rode, M.; Borchardt, D. Carbon and nutrient export regimes from headwater catchments to downstream reaches. *Biogeoosciences* 2017, 14, 4391–4407. [CrossRef]


86. Darmody, R.G.; Ruiz Diaz, D. *Dredged Sediment: Application as an Agricultural Amendment on Sandy Soils*; Illinois Sustainable Technology Center: Champaign, IL, USA, 2017.