Rothamsted Repository Download

A - Papers appearing in refereed journals

Granger, S. J., Harris, P., Upadhayay, H., Sint, H. M., Pulley, S., Stone, M., Krishnappan, B. G. and Collins, A. L. 2021. Novel approaches to investigating spatial variability in channel bank total phosphorus at the catchment scale. *Catena.* 202 (article), p. 105223. https://doi.org/10.1016/j.catena.2021.105223

Catena 202 (2021) 105223

Contents lists available at ScienceDirect

Catena

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

Novel approaches to investigating spatial variability in channel bank total phosphorus at the catchment scale

S.J. Granger[a,*], P. Harris[a], H.R Upadhayay[a], H. Sint[a], S. Pulley[a], M. Stone[b], B.G. Krishnappan[c], A.L. Collins[a]

[a] Rothamsted Research, North Wyke, Okehampton, Devon EX20 2SB, United Kingdom
[b] Department of Geography and Environmental Management, University of Waterloo, Waterloo N2L 3G1, Canada
[c] National Water Research Institute, Environment Canada 867 Lakeshore Road, Burlington, Ontario L7R 4A6, Canada

ARTICLE INFO

Keywords:
Riverbank sediment
Stream order
Land use
Spatial
Sample design
Geographically weighted

ABSTRACT

Phosphorus (P) is often a limiting nutrient that leads to the eutrophication of aquatic systems. While dissolved P forms are the most bioavailable, the form, mobility, transport and fate of P are directly related to its association with fine-grained riverine sediment. Therefore, to implement successful P catchment management strategies it is important to understand the relative contribution of different sediment sources to P loads across the river continuum. While agricultural topsoil and, to a lesser extent, riverbed sediment are important sources of sediment-associated P, channel banks have been shown to be an important sediment source in some catchments. However, comparatively little is known about the P concentration and corresponding spatial variability in channel bank sediment and the associated implications for catchment management. The present study examines the spatial variability of P associated with channel bank profiles within a series of three nested catchments using both non-spatial and spatial statistical methods, where for the latter, a novel spatial approach was used to estimate the spatial averages and variances of P in channel bank sediment along the stream network. Channel bank P concentrations were compared to factors such as catchment scale, stream order, land use, bank exposure and location along the stream network. Concentrations of TP ranged between 129.6 and 1206.9 mg P kg^{-1} of which the water extractable P (WEP) content ranged from 0.01 to 0.12%. Stream order was found to influence TP concentrations, while land use and catchment scale provided only a moderate influence. This suggested that focussing channel bank sampling strategies at the largest catchment scale would capture key drivers of TP variability provided stream order is sufficiently represented. Whether the bank was had limited vegetation and was exposed and potentially eroding had a slight influence on TP variability in second-order stream banks in the larger of the two nested catchments. However, the slightly lower TP concentrations measured at these sites indicated that banks that are actually eroding may be contributing less TP than the total channel bank TP values measured across the catchments as a whole. The results of an explicitly spatial analysis demonstrated that local channel bank TP averages and TP variances vary along the stream network. Specifically, the most accurate spatial predictor of TP was local TP means with the use of 'crow flies' rather than stream network distances. Local TP variances were used to provide optimal designs for future channel bank TP sampling campaigns, given available resources. Throughout, both standard and outlier-resistant statistical analyses were applied to improve interpretation of the study findings.

1. Introduction

There is a growing requirement to improve catchment management strategies that control phosphorus (P) mobilisation and delivery to maintain water quality and ecological status of watercourses globally (McDowell et al., 2016). Surface waters are particularly sensitive to inputs of P because relatively low concentrations of bioavailable P, an order of magnitude lower than soil available P concentrations required for crop growth (Heathwaite and Dils, 2000), can accelerate the growth of nuisance algae and impair water quality (Carpenter et al., 1998; Conley et al., 2009). Recognizing that P represents a key pressure on the ecological status of rivers, many land use policies require implementation of management programmes to reduce P input to surface waters in order to improve water quality (EEC, 2000). Whilst soluble reactive P

* Corresponding author.

https://doi.org/10.1016/j.catena.2021.105223
Received 28 May 2020; Received in revised form 9 December 2020; Accepted 30 January 2021
Available online 26 February 2021

(SRP) is primarily targeted by such legislation, in most rivers the majority of P instead enters watercourses in association with fine sediments (Lloyd et al., 2019; Meybeck and Helmer, 1989; Royer et al., 2006). Fine grained sediments have a high surface area per unit mass and contain metal-oxy hydroxide surface coatings that act as sources for bioavailable SRP due to P exchange with waters (Stone and English, 1993; Stone and Mudroch, 1989; Taylor and Kunishi, 1971). Agricultural land has been identified as one of the most important diffuse sources of P delivery to freshwater systems. Morse et al. (1993) estimated that 45% of UK P inputs to rivers originated from agriculture, while Zhang et al. (2014) estimated that the contribution of total P (TP) to rivers from agriculture across England and Wales was 31%. Therefore, to reduce the impact of P in aquatic environments, P inputs from different sources should be quantified (Heathwaite et al., 2005; Kronvang et al., 2007) so that mitigation options can be most efficiently targeted (Haygarth et al., 2009). This is also important to address the potential for spatial mismatches between the expected efficacy of on-farm environmental measures and the observed reductions in pollutant loads within catchments (e.g. Biddulph et al., 2017).

Implementation of successful P catchment management strategies requires knowledge of all potential P sources in the catchment(s) and to understand their relative contribution, particularly those associated with sediment. This has been achieved in some cases through sediment source fingerprinting techniques using tracers and statistical un-mixing models (Collins et al., 2017; Haddadchi et al., 2013). Using these tracing approaches within rural environments, identifiable sediment sources typically include pasture and arable land, un-metalled roads and metalled road damaged verges (Collins et al., 2012; Minella et al., 2009), and eroding channel banks (Collins et al., 1997; Owens et al., 2000; Slattery et al., 2000; Walling et al., 1999).

While the erosion of agricultural top soils is generally the most important source of suspended sediment, channel banks have also been shown to act as an important contributor to sediment loads in many rivers (e.g., Evans et al., 2003; Imeson et al., 1984; Laubel et al., 2003; Rode et al., 2018; Zaimes et al., 2008a). Information on the TP content in channel bank sediment is limited and even less information on the water extractable P (WEP) concentrations is available (Fox et al., 2016). However, some authors report sediment-associated P loadings derived from channel banks as a major contributor of P loads in some rivers. Sekely et al. (2002) estimated that in the Blue Earth River in Minnesota US, 7–10% of TP load originated from channel bank erosion, while Roseboom (1987) reported that up to 56% of the P load in central Illinois originated from channel banks in floodplains. In some Danish streams, channel bank erosion can supply up to 90% of P loads (Kronvang et al., 1997). Working in 12 UK sub-catchments, Walling et al. (2008) reported that channel and sub-surface sources accounted for up to 55% of the sampled suspended sediment load which, in turn, accounted for up to 43% of the sediment-associated P flux. Despite this, most work on P inputs has left the contribution of channel banks largely unreported (Fox et al., 2016). Where it is reported, studies tend to rely on the integration of resource demanding source fingerprinting procedures and information on the typical (e.g. average) TP concentrations in the sampled source materials including eroding channel banks (Walling et al., 2008). Therefore, to derive a meaningful estimate of TP concentrations, systematic sampling of channel bank sources is required. However, important research questions related to channel bank TP concentrations include, what are the TP concentrations in channel banks, how do they vary spatially (Fox et al., 2016), and in the case of larger catchments, how can streambanks as a potential P source, be better understood so that their contribution can be reliably accounted for? To this end, a novel study was undertaken within a rural UK catchment to examine the magnitude and distribution of channel bank TP concentrations, and the implications for future sampling. Given the research gap identified for WEP concentrations in channel banks (Fox et al., 2016), this work also compared WEP against TP concentrations. Comparisons of WEP to iron (Fe) and aluminium (Al) content were also undertaken as the amount of WEP that can be released is heavily dependent on the soil components that sequester and retain P, the most significant typically considered to be Fe and Al (Nair, 2014).

Specifically, our study rigorously quantified channel bank TP concentrations within a series of three nested catchments using both non-spatial and spatial statistical methods. We investigated a dataset of $n = 60$ samples and evaluated changes in TP distributions according to: (i) Catchment (not size but its 'nested' position); (ii) Stream Order; (iii) Land Use; (iv) degree of channel bank 'exposure', and; (v) spatial location along the stream network. The TP investigations consisted of non-spatial methods, such as ANOVAs and regressions for parts (i) to (iv), above. Investigations along the stream network for part (v), above, were explicitly spatial, wherein local averages and variances were found within a geographically weighted modelling framework (Brunsdon et al., 2002; Harris and Brunsdon, 2010; Harris et al., 2014b; Gollini et al., 2015).

Our implementation of localized summaries to a dataset collected over a stream network was entirely novel, where both Euclidean ('as the crow flies') and stream network distances were investigated; complementing existing applications of local regression with road network and travel time distance metrics in an urban context (Lu et al., 2014). The use of local statistics also represent a simple alternative to existing stream network methodologies stemming from different statistical paradigms, which commonly refer to the investigation of water chemistry in the stream itself, where flow direction is accounted for (e.g., Gallacher et al., 2017; Ver Hoef et al., 2006; Ver Hoef and Peterson, 2010). A further advance was demonstrated in that optimal designs for future TP sampling campaigns were found using the local variances as inputs (following Harris et al., 2014b), again using different distance metrics. Throughout our work, both standard and robust (outlier-resistant) analyses were undertaken in parallel, where the latter mitigates against poor estimation and inference due to outlying TP concentrations.

2. Methodology

2.1. Study site and design

The study was conducted in a largely rural part of west Devonshire in the UK within the headwaters of the River Taw (Fig. 1.). The upper reaches of the River Taw have been assessed in conjunction with the chemical requirements of the EU Water Framework Directive (WFD) and judged to be failing as a result of elevated SRP levels and as failing to meet 'good status' due to the impacts of excessive sediment inputs on fish, including lithophilous species, such as salmonids, dependent on clean riverbed gravels for key early life stages (Finn, 2007). The river network that comprises the Upper River Taw catchment is constrained within a landscape of rolling hills which tend to lead to deeply incised channels with very limited floodplain zones which are seldom inundated. The channels in the south of the catchment, where the River Taw rises from an upland plateau, occur on low agricultural intensity extensive grasslands with limited tree cover. The river runs through a mosaic of habitats including blanket bog, upland heath, and acid grassland typically vegetated with grasses (e.g. Agrostis capillaris and Festuca ovina), shrubs (e.g. Calluna vulgaris and Ulex gallii), herbs (e.g. Galium saxatile and Potentilla erecta) and bracken (Pteridium aquilinum). Here a very low livestock stocking density means channel banks are untrampled to any significant degree and no significant channel manipulation has occurred for several hundred years. Once the River Taw leaves this area the land use become significantly more agricultural and is predominantly grassland, with only a limited amount or arable cropping. The nature of the channel system in this lower reach still tends to be unmanipulated, but the larger channels are typically fenced off to exclude animal access and are often lined with large (e.g. Quercus robur and Fraxinus excelsior) and small (e.g. Alnus glutinosa and Crataegus monogyna) tree species and with good vegetation cover from shrubs (e. g. Rubus fruticosus and Hedera helix), grasses (e.g. Deschampsia

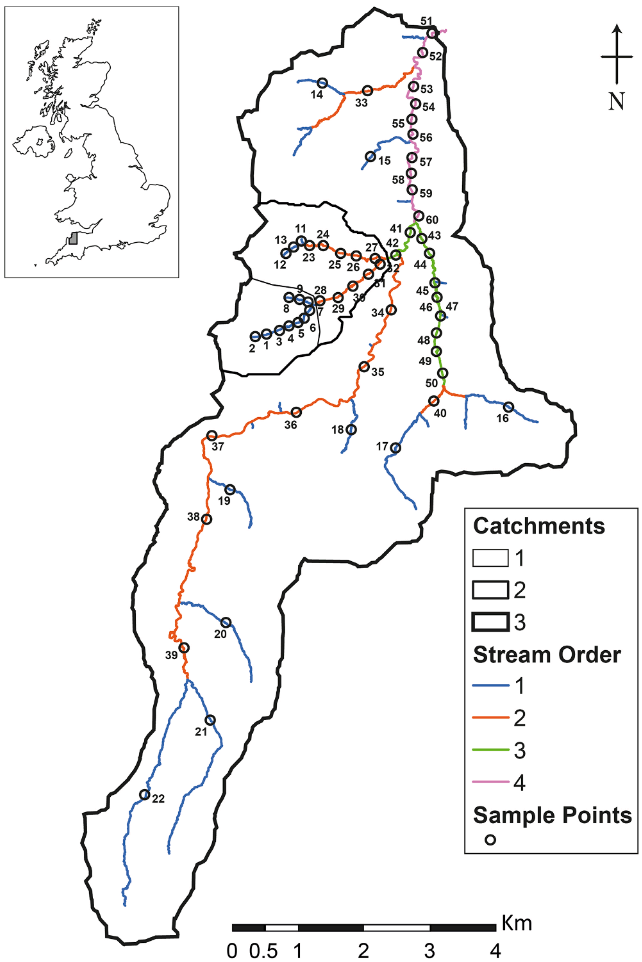

Fig. 1. Study nested Catchments (1, 2 and 3) shown with corresponding channel bank sampling site numbers (1–60) and Stream Order numbers (1–4).

flexuosa and Poa nemoralis), sedges (e.g. Carex pendula and Carex remota) and herbs (e.g. Hyacinthoides non-scripta and Filipendula ulmaria). The smaller feeder streams can be open to animal access in some cases although they do not tend to be heavily trampled and are still in places fenced off. Where arable cropping occurs, these tend to sit within grassland riparian zones and cropping is not undertaken close to the channel bank itself. Examples of typical channel reaches are contained in Supplementary Information (Figures S1a-f). Within the headwaters of the River Taw catchment, an instrumented landscape observatory has been developed to monitor various physio-chemical parameters and river discharge in conjunction with a research programme examining on-farm best management pathways for sustainable agriculture. In this observatory, three monitored nested catchments were used to define the different scales of the sampling locations in the study area. These catchments were 1.7, 4.4 and 41.3 km^2 in area and are shown in Fig. 1, labelled as catchments 1 (smallest), 2 and 3 (largest), respectively. Recognising that as scale increases, so the nature of the hydrological network changes, channel bank samples were collected based on stream order within each catchment. The ArcGIS (v10.4) - Spatial Analyst Hydrology toolset was used to simplify the classification and provide clarity and consistency. A linear stream feature layer was created using OS MasterMap (http://edina.ac.uk/digimap) which, in turn, was 'burned' into the Tellus South West digital elevation model (DEM) (http://www.tellusgb.ac.uk/) to force the model to follow, as closely as possible, the actual drainage network. The resulting DEM was used to create a Flow Direction and Accumulation grid layer with a reasonable size stream grid mirroring that of the actual stream network.

The ArcGIS Stream Order tool was used to assign the layer with stream order values which ranged from 1 to 4 across the entire study area, with a single incidence of stream order 4 ending at the headwater.

Within each of the three catchments, 10 sampling sites were randomly ascribed per stream order, with a number of these sites being included in each of the successive larger catchment within which the nested catchments occurred. Within Catchment 1, all channels were classified as being 1st stream order. Within Catchment 2, channels were classified as both 1st and 2nd order of which 7 of the 1st order sampling sites correspond to those in Catchment 1. Within Catchment 3, channels were classified as being between 1st and 4th order, of which 1 sampling site corresponded to that within Catchments 1 and 2, with the remaining 9 occurring on new 1st order channels. Further, two 2nd order channel sampling sites correspond to those within Catchment 2, with 8 new sampling locations, and all the 3rd and 4th order channel sampling sites exist solely within Catchment 3. This sampling strategy resulted in $n =$ 60 samples which are depicted in Fig. 1 and details contained within Table 1. The sampling strategy resulted in ten measurements which were effectively representative of multiple sites to form an $n = 70$ rather than an $n = 60$ dataset. Section 2.3.4 details how a judicious sub-setting of the $n = 70$ dataset ensures statistically unbiased analyses.

2.2. Sample collection, preparation and analysis

At each sample location, a composite sample of channel bank material was collected. The channel bank was visually assessed, and samples collected from a section of channel which had a similar profile on each bank. Samples were collected from the 'bankfull' channel profile (Rosgen, 1994) from one of the banks at each sampling location. Specifically, five (3 cm deep by 5 cm diameter) bank cores were collected and bulked into one composite sample, equidistantly vertically down the channel bank between the highest point of the 'bankfull' channel profile down to the occurrence of channel bed sediment, or to where bedrock started to occur. Recognising that different site specific factors might influence bank P content, at each site the predominant land use above the sampled bank was noted and also whether the bank had <50% vegetation where sampled (hereafter referred to as 'exposed') to differentiate between banks which are potentially actively eroding from those which might be considered more stable.

Cores were dried at 105 °C for 24 h before being passed through a 2 mm sieve to remove large stones and organic debris. Post thorough mixing, a sub-sample of the <2 mm material was then milled to a fine powder before being passed through a 63 μm sieve. The <63 μm material was then analysed for both TP and WEP, and also total Fe and Al. Total P, Fe and Al were determined following an aqua regia (hydrochloric acid: nitric acid; 80:20 V/V) digestion in open tube digestion blocks. The acids were removed by volatilisation and the residue dissolved in nitric acid (5% V/V) and filtered through a Whatman 40 filter paper. The resultant solution was then analysed for TP, Fe and Al on an Optima 7300 DV Inductively Coupled Plasma - Optical Emission Spectrometer (ICP-OES), with analytical limits of detection (LOD) of 12.0, 100.1 and 78.4 mg kg^{-1}, respectively. Analytical performance was monitored using sample digestion, and analytical standards and replicates. Water extractable P was determined by shaking 20 g of <63 μm soil in 80 ml of deionised water (volume adjusted for soil mass to ensure constant ratio of material) for 1 h at 25 °C. The solutions were then centrifuged at 16.9*g* for 10 min and filtered using Whatman no. 42 filter papers. Concentrations of orthophosphate were determined in the resultant solutions according to Murphy and Riley (1962) using a discrete photometric analyser (Thermo-Fisher Aquakem 250, Loughborough, UK) with a LOD of 1.5 mg P l^{-1}.

Table 1
Sample site numbers (see Fig. 1) shown with their Catchment (C1, C2 and C3), Stream Order (SO1 to SO4), TP (mg kg^{-1}) and WEP (mg kg^{-1}) concentrations. The ten measurements that effectively represent multiple sites are highlighted, as these form the $n = 70$ data set, where measurements at site # 1, 4, 6, 8, 9, 10, 25 and 29 are repeated once (highlighted in blue), while the measurement at site # 3 is repeated twice (highlighted in orange).

Site #	C1	C2	C3	SO	TP	WEP	Site #	C1	C2	C3	SO	TP	WEP
1	X	X		1	1060	1.06	**31**		X		2	450	0.09
2	X			1	343	0.27	**32**		X		2	461	0.08
3	X	X	X	1	799	0.94	**33**			X	2	570	0.08
4	X	X		1	703	0.26	**34**			X	2	845	0.31
5	X			1	543	0.03	**35**			X	2	749	0.14
6	X	X		1	754	0.43	**36**			X	2	658	0.08
7	X			1	699	0.10	**37**			X	2	405	0.08
8	X	X		1	1083	0.33	**38**			X	2	510	0.08
9	X	X		1	714	0.45	**39**			X	2	816	0.14
10	X	X		1	747	0.09	**40**			X	2	1131	0.09
11		X		1	409	0.26	**41**			X	3	684	0.09
12		X		1	380	0.07	**42**			X	3	631	0.16
13		X		1	130	0.05	**43**			X	3	974	0.20
14			X	1	357	0.17	**44**			X	3	996	0.13
15			X	1	489	0.22	**45**			X	3	1207	0.15
16			X	1	357	0.15	**46**			X	3	928	0.11
17			X	1	950	0.69	**47**			X	3	983	0.14
18			X	1	546	0.06	**48**			X	3	1064	0.09
19			X	1	398	0.09	**49**			X	3	945	0.07
20			X	1	691	0.04	**50**			X	3	1103	0.14
21			X	1	629	0.04	**51**			X	4	725	0.18
22			X	1	774	0.27	**52**			X	4	727	0.21
23		X		2	600	0.07	**53**			X	4	973	0.85
24		X		2	496	0.04	**54**			X	4	895	0.10
25		X	X	2	443	0.06	**55**			X	4	586	0.07
26		X		2	558	0.17	**56**			X	4	592	0.17
27		X		2	532	0.09	**57**			X	4	861	0.14
28		X		2	796	0.24	**58**			X	4	924	0.23
29		X	X	2	417	0.08	**59**			X	4	972	0.22
30		X		2	489	0.08	**60**			X	4	876	0.29

2.3. Statistical methods

2.3.1. Analyses of channel bank TP concentrations with respect to key drivers

Major drivers of TP variability in the sampled channel bank materials were evaluated statistically using a series of conditional analyses according to Catchment, Stream Order, Land use (on the sampled bank) and degree of bank exposure (termed 'Exposed'). Conditional density plots and boxplots are presented to visualise differences among the conditional TP distributions. Formal statistical analyses were conducted, using ANOVAs and Kruskal-Wallis (KW) rank sum tests (Vargha and Delaney, 1998) in order to test whether observations of TP variability in channel banks due to the informal conditional analyses above, were statistically significant. Where appropriate, the ANOVAs and KW tests specific to the Catchment, Stream Order or Land Use were supplemented by a respective post-hoc analysis (Tukey Honest Significant Differences (HSD) (Tukey, 1949) and the Dunn test (Dunn, 1964) in order to determine *which* Catchment or *which* Stream Order or *which* Land Use had significantly different TP distributions (as an ANOVA or KW test only indicates at least one category is different, but not *which* category). Further statistical inferences were conducted via regressions, wherein TP was predicted using Stream Order, Land Use, and Exposed as categorical predictors of TP in order to further assess their significance in discriminating different sources of variation in the channel bank sample TP concentrations at different catchment scales.

2.3.2. Spatial analysis of channel bank TP concentrations using different distance metrics

A non-stationary spatial analysis was conducted wherein local TP means, medians, standard deviations (SDs) and median absolute deviations (MADs) were calculated at a pre-defined set of regular points ($n = 110$) along the stream network, informed by measured TP concentrations (i.e. the $n = 60$ sample dataset was used to calculate the $n = 110$ dataset). This was implemented to provide insight into how TP concentrations vary along the stream network with respect to spatial changes in channel bank TP averages (means and medians), but also with respect to spatial changes in TP variability (SDs and MADs). Our main analytical goal was to investigate spatial changes in TP variability rather than TP averages, but both are presented in recognition that they are intrinsically linked, as described below.

To calculate the local summary statistics, a geographically weighted moving-window approach was used where the weighting was determined by a bi-square distance-decay kernel function (Gollini et al., 2015) using both Euclidean ('as the crow flies') and stream network distances. The size of the moving-window (the kernel bandwidth) was taken as the number of nearest TP measurements (i.e. a subset of the $n = 60$ sampled points, from which the local statistic was calculated) to the calculation point (i.e. an unsampled point, of which there are $n = 110$). Bandwidth selection is critical to any localized model calibration and was determined optimally via leave-one-out cross-validation (Brunsdon et al., 1996; Harris et al., 2014b), but for the local TP means and medians, only. The same bandwidths were respectively assumed optimal for use with the local TP SDs and local TP MADs. This strategy for bandwidth choice was considered reasonable, given that averages commonly scale directly to variances for many environmental processes, and given that no ideal cross-validation approach exists for finding an optimal bandwidth for local measures of variability (Harris et al., 2014b). Observe that local means and local medians are both spatial interpolators (predictors), and as such, the cross-validation exercise directly provides a measure of their overall prediction accuracy. Here, the cross-validation score is specified as the residual sum of squares based on leave-one-out predictions (Brunsdon et al., 1996), and the lowest cross-validation score determines the optimal bandwidth. Mathematical formulae using Euclidean distances for local means, medians and SDs are given in Brunsdon et al. (2002), while those for local MADs are given in Harris et al. (2014b). Extensions to cater for stream network distances simply consists of substituting the associated distance matrices into the respective local statistic formulae.

Monte Carlo randomization tests were conducted to identify locations along the stream network where the given local TP summary statistic was 'significantly' different from such a local statistic found by chance or artifacts of random variation in the data. The Monte Carlo tests were only conducted at measured TP locations, for local averages only, and were reported at the 95% level. The Monte Carlo tests follow those described in Harris and Brunsdon (2010) but now using local averages calculated with different distance metrics. As a final demonstration of the local modelling toolkit, optimal sample designs were found for future TP sampling campaigns using the local variances as inputs to a location-allocation algorithm (Teitz and Bart, 1968), again using different distance metrics, extending techniques given in Harris et al. (2014b) in a lake freshwater acidification context and those given in Kanaroglou et al. (2005) in an urban air pollution context.

2.3.3. *Use of robust statistics and models*

Both standard and robust statistics/models were used for an increased likelihood of outputs being compromised by outlying TP concentrations due to low sample numbers (see Section 2.3.4). Specifically, the median and MAD represent robust alternatives to the mean and SD, respectively; the KW rank sum test represents a robust alternative to an ANOVA; and a least trimmed squares (LTS) regression (Rousseeuw and Leroy, 1987) was fitted to represent a robust alternative to an ordinary least squares (OLS) regression fit. Observe that MAD is preferred to the inter quartile range (IQR) or the Qn scale estimator as our chosen robust variance statistic. This is because the IQR is a rather crude statistic, while the Qn scale estimator (Rousseeuw and Croux, 1993) does not depend on an estimated average (i.e. the median in this case), whereas MAD does. For our approach, the dependence of the TP variance estimate to its corresponding average is considered important given the optimal local variance bandwidth is assumed the same as the optimal local average bandwidth, from the spatial analysis above. Thus, MAD is used, even though the Qn scale estimator is a more efficient statistic (Rousseeuw and Croux, 1993).

To help facilitate standard to robust comparisons, statistical tests associated with the methods of Section 2.3.1 were reported at the 90% level of confidence, so that small but critical changes in significance could be observed. Tests associated with the methods of Section 2.3.2 were at the 95% level, as already stated.

2.3.4. *Addressing potential sampling bias*

The next consideration was one of an inherent bias due to the bank material sampling strategy, which focused on Stream Order coverage, rather than any other consideration, and was itself limiting due to available resources. Thus, because the strategy was to sample evenly for each Stream Order, it was known from the outset that some regions of the stream network would be sampled more intensively than others (Fig. 1). Furthermore, the nested nature of the three catchments allowed ten measurements to be representative of multiple sites, meaning that $n = 10$ sites were in Catchment 1, all of Stream Order 1; $n = 20$ sites were in Catchment 2, with sites at Stream Orders 1 and 2, both of size $n = 10$; and $n = 40$ sites were in Catchment 3, with sites at Stream Orders 1, 2, 3 and 4, all of size $n = 10$. This gives rise to the $n = 70$ dataset (Table 1).

Thus, to mitigate against likely biases of over- and under-representation, all non-spatial analyses (summary statistics, plots, ANOVAs, regressions, etc.) were applied to the $n = 70$ dataset but almost exclusively on a catchment by catchment basis as illustrated in Fig. 2. This approach ensured that Stream Order sample size remains consistent (i.e., to datasets of $n = 10$, $n = 20$ and $n = 40$, respectively). The single exception to this approach was when TP distributions were assessed for differences across the three catchments, and here the $n = 70$ dataset was used directly but noting potential bias of unbalanced information across catchments. Effects of sampling bias on the spatial analyses using the $n = 60$ dataset are discussed in Section 4.

2.3.5. *Analysis of channel bank WEP concentrations*

The WEP concentrations were subjected to a standard, linear correlation analysis only, where the aim was to assess, statistically, the strength of WEP relationships to the TP content of the channel bank material and to the corresponding combined Fe and Al content (termed 'Fe + Al'). Corresponding tests were reported at the 95% level.

3. Results

3.1. *Overview*

Measured TP concentrations in channel bank sediment ranged

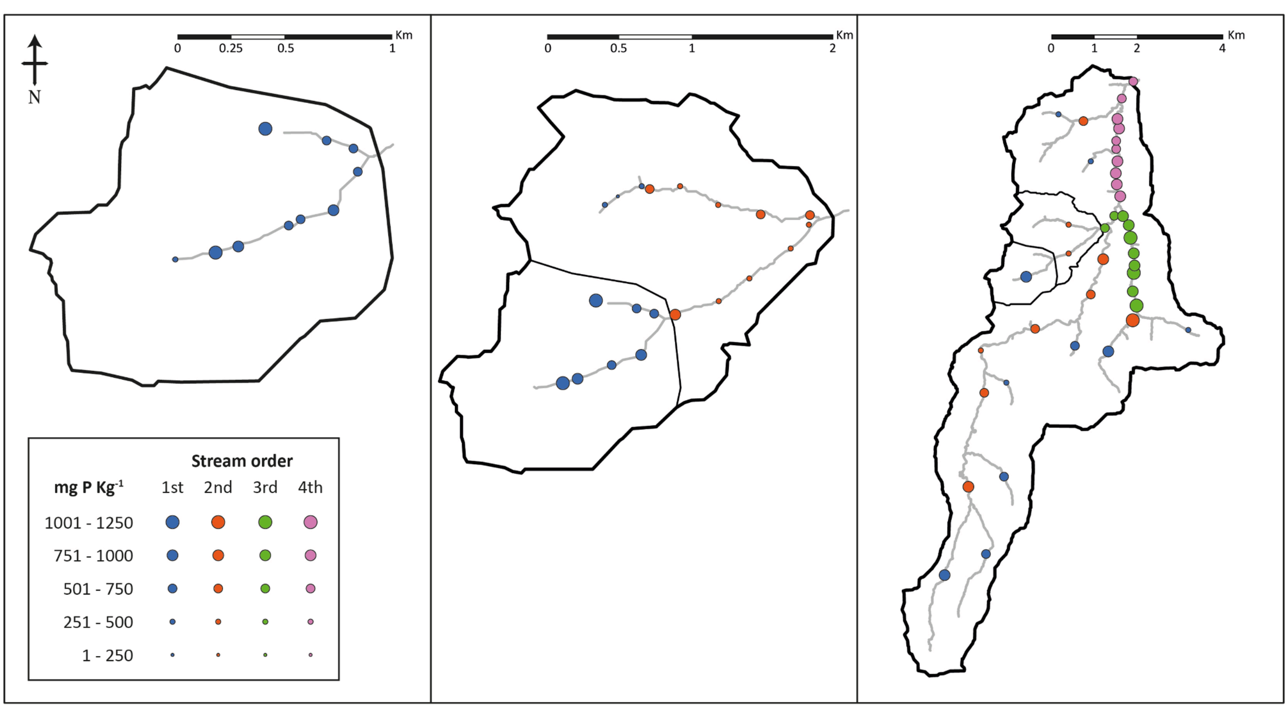

Fig. 2. Spatial distribution of channel bank material TP (mg kg^{-1}) by Catchments 1, 2, and 3; and by Stream Order; with sample sizes of n = 10, 20 and 40, respectively.

between 129.6 and 1206.9 mg P kg^{-1}. The WEP concentrations ranged between 0.03 and 1.06 mg P kg^{-1} (Table 1) which represented between 0.01 and 0.12% of the TP. Mapped summaries of TP concentrations by Catchment and Stream Order are presented in Fig. 2. Through density and boxplots, Fig. 3 displays TP concentrations conditional to Catchment and to Stream Order, respectively. Fig. 4 provides boxplots of TP concentrations conditional to Land Use and to Exposed, respectively.

Statistical analyses for the significance of the same conditional TP distributions are presented in Table 2 for ANOVA / K-W tests. Regressions are given in Table 3, where significant predictors of TP variability (i.e., from Stream Order, Land Use and Exposed) are highlighted. Results of the spatial analyses are given in Figs. 5 and 6, where local TP averages and local TP variances are mapped to show how these statistical summaries are not constant, but instead, vary along the sampled stream network in the upper River Taw catchment. It is useful to present the results in terms of TP concentrations conditional to: (a) Catchment, (b) Stream Order, (c) Land Use, (d) Exposed and (e) spatial location along the sampled stream network.

3.2. Total P concentrations by catchment

Measured TP concentrations in channel bank samples for Catchment 2 were significantly different to those found in Catchment 3 but not significantly different to those in Catchment 1, and TP concentrations for Catchment 1 were not significantly different to those in Catchment 3 (Table 2). The level of significance tends to be at the 90% level rather than the 95% level, especially when considering robust methods (*p*-values for KW rank sum and Dunn test were 0.072 and 0.074, respectively). The TP distributions in each catchment strongly overlap (Fig. 3A and C). The TP distribution for Catchment 1 is multi-modal (which is caveated due to its small sample size of only $n = 10$). Overall, there is only moderate evidence for differences in TP variability due to Catchment. These results were based on the $n = 70$ dataset and subsets thereof.

3.3. Total P concentrations by stream order

Using the $n = 40$ dataset associated with Catchment 3 only, channel bank sample TP concentrations for Stream Order 1 were significantly different than in Stream Order 3 (at the 95% level) and Stream Order 4 (at the 90% level, only), and TP concentrations for Stream Order 2 were significantly different than in Stream Order 3 (at the 95% level) (Table 2). The TP distributions by each stream order overlap each other as depicted in the conditional densities (Fig. 3B) and conditional boxplots (Fig. 3D), where the TP distribution for Stream Order 3 is multi-modal. Higher TP concentrations tended to be in the third and fourth stream orders, while lower TP concentrations tended to be in first and second stream orders. Regression analyses (Table 3) indicate that Stream Orders 3 and 4 are both significant predictors/drivers of TP variability, where the robust regression provides stronger evidence of this (noting that Stream Order 1 drops out of the regression analysis, in all instances).

Conversely, no significant differences were observed between Stream Orders 1 and 2 upon reviewing the results given in Table 2, for the $n = 20$ dataset associated with Catchment 2 (noting that only these two Stream Orders are present). Although, in the corresponding regression analysis in Table 3, Stream Order is a significant driver of TP variability in the sampled channel bank materials for the robust regression fit only. Overall however, the results change little depending on whether a standard or robust model is used, suggesting minimal influence of outlying TP concentrations with respect to Stream Order. In summary, there is moderate to strong evidence for differences in TP variability due to Stream Order.

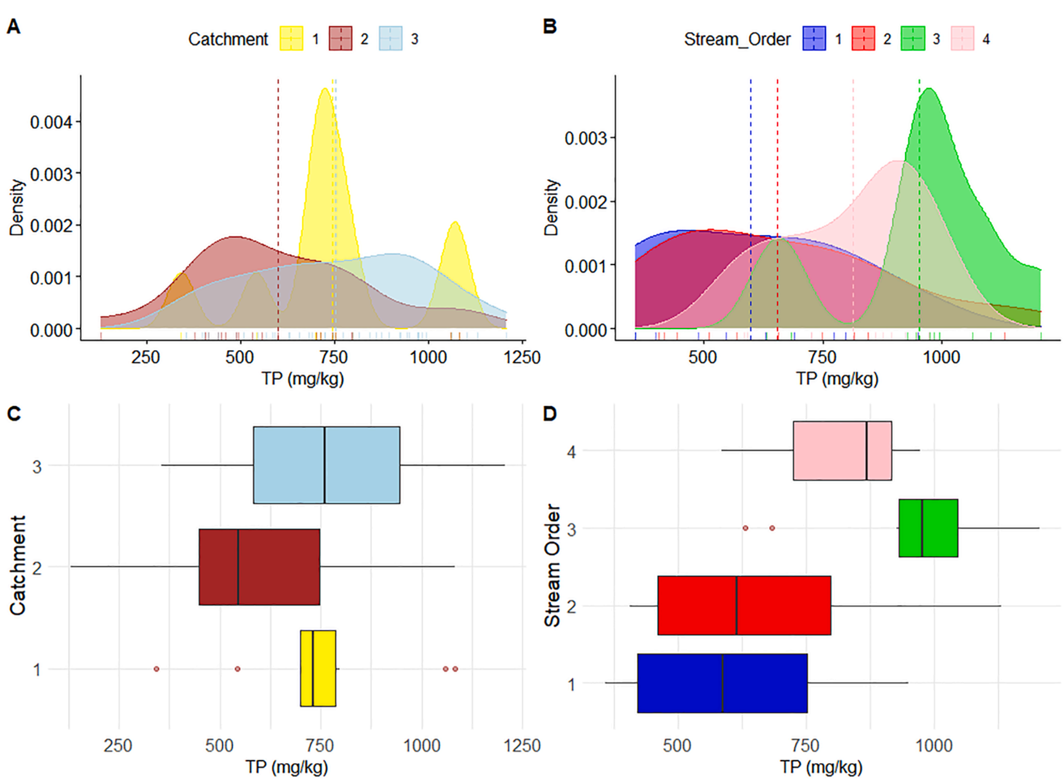

Fig. 3. (A) Density plots for TP (mg kg^{-1}) by Catchment with conditional means highlighted (vertical lines) at 744.3, 600.9 and 754.5 mg kg^{-1} (for $n = 70$ dataset); (B) Density plots for TP by Stream Order (Catchment 3 only, $n = 40$ dataset) with conditional means highlighted (vertical lines) at 398.9, 654.4, 951.6 and 813.1 mg kg^{-1}; (C) Boxplots for TP by Catchment (for $n = 70$ dataset); (D) Boxplots for TP by Stream Order (Catchment 3 only, $n = 40$ dataset).

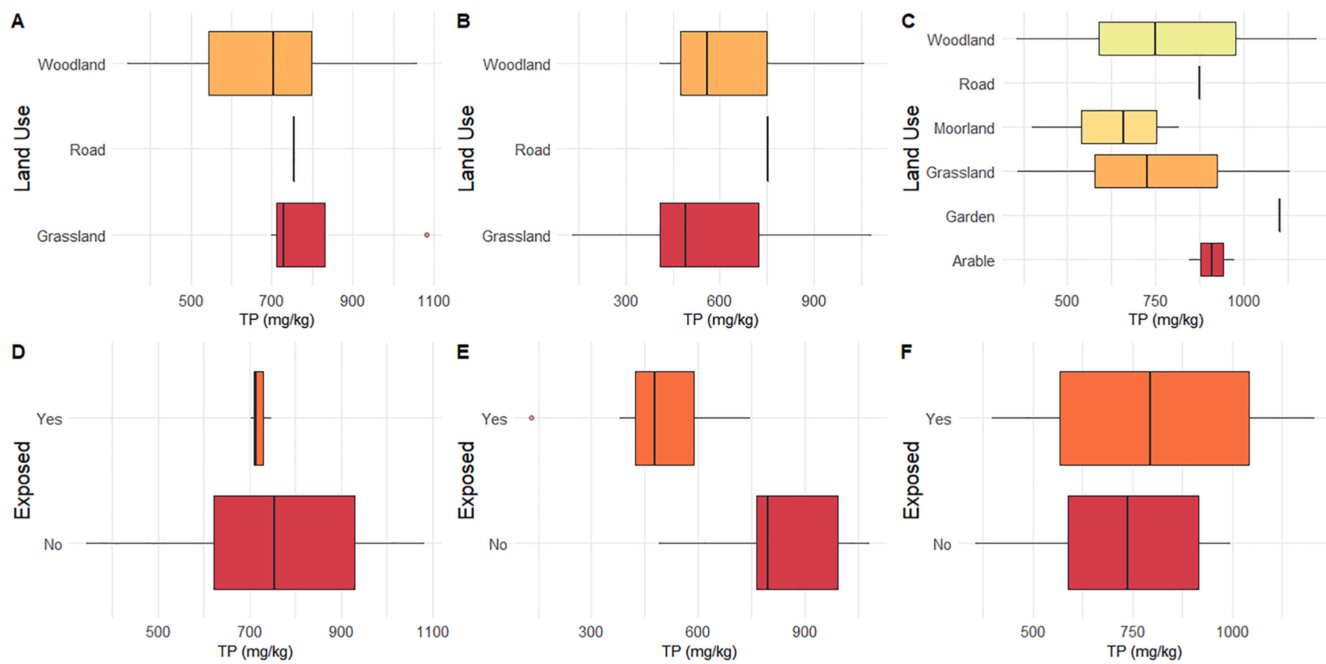

Fig. 4. Boxplots for TP (mg kg^{-1}) by Land Use for: (A) Catchment 1, (B) Catchment 2, and (C) Catchment 3. Sample sizes of n = 10, 20 and 40, respectively. Boxplots for TP by Exposed for: (D) Catchment 1, (E) Catchment 2, and (F) Catchment 3. Sample sizes of n = 10, 20 and 40, respectively.

Table 2
Significance tests at 90% level for factors driving TP variability in channel bank samples: Catchment, Stream Order, Land Use and Exposed. Tests used: ANOVA, Tukey Honest Significant Differences (HSD) (post-hoc test for ANOVA), Kruskal-Wallis (KW) rank sum test and Dunn test (post-hoc test for KW rank sum).

Test	Catchment (and sample size)	ANOVA	Tukey HSD	KW rank sum	Dunn
Catchment (C1, C2 and C3)	NA (n = 70)	0.052	0.045 (C2 with C3)	0.072	0.074 (C2 with C3)
Stream Order (SO1-2)	C2 (n = 20)	Not significant	–	Not significant	–
Stream Order (SO1-4)	C3 (n = 40)	0.001	0.001 (SO1 with SO3) 0.007 (SO2 with SO3) 0.079 (SO1 with SO4) Rest not significant	0.002	0.003 (SO1 with SO3) 0.009 (SO2 with SO3) 0.089 (SO1 with SO4) Rest not significant
Land Use	C1 (n = 10)	Not significant	–	Not significant	–
Land Use	C2 (n = 20)	Not significant	–	Not significant	–
Land Use	C3 (n = 40)	Not significant	–	Not significant	–
Exposed	C1 (n = 10)	Not significant	–	Not significant	–
Exposed	C2 (n = 20)	0.001	–	0.004	–
Exposed	C3 (n = 40)	Not significant	–	Not significant	–

3.4. Total P concentrations by land use

Land Use does not significantly influence TP variability in the channel bank samples (Table 2 for the n = 10, n = 20 and n = 40 datasets, associated with Catchments 1, 2 and 3, respectively; see also Fig. 4). Although, in the corresponding regression analysis in Table 3, Grassland and Woodland (for the n = 40 dataset) were significant drivers of TP variability, with negative associations (coefficient signs), for the robust regression fit only.

3.5. Total P concentrations by 'Exposed'

The degree of bank exposure significantly influenced TP variability in channel bank sediments in Catchment 2, only (Tables 2 and 3; see also Fig. 4). Test results are effectively no different between standard and robust models, suggesting minimal influence of outliers with respect to the TP and Exposed relationship. Investigations (not shown) to assess whether TP concentration depended on both Exposed and Stream Order found that no clear discernable patterns were observed (this is returned to in the discussion section below).

3.6. Total P concentrations by stream network location

The results of the spatial analysis are given in Figs. 5 and 6, where local TP means, medians, SDs and MADs were calculated along the network at a pre-defined set of regular points using both 'crow flies' and stream distances. Results are dependent on the specified bandwidths, which were found optimally at 58 and 58 nearest neighbours for TP means with 'crow flies' and stream distances, respectively; and 58 and 33 nearest neighbours for TP medians, also with 'crow flies' and stream distances, respectively. Corresponding cross-validation scores were 2473, 2498, 10,457 and 43,651, respectively. Thus, for a local TP mean using 'crow flies' distances, the nearest 58 of a possible 60 TP measurements were used in its calculation, but where the nearest measurements had a greater influence than those further away. Alternatively, for a local TP median using stream distances, the nearest 33 TP measurements were used in its calculation, again where measurements most near had the greatest influence. Reporting bandwidth results provides important insight into the spatial process (Harris and Brunsdon, 2010; Harris et al., 2014b), where for TP its local mean (with its relatively large bandwidth) will tend to its global value, while its local median (with its relatively small bandwidth) will not. Similar

Table 3
Regression fits. Only p-values < 0.1 shown (i.e., 90% significance level). All variance inflation factors (VIFs) < 2.4 (for gauging predictor variable collinearity).

	Catchment 2 Model only (n = 20)		Catchment 3 Model only (n = 40)	
OLS Regression Coefficient	**Estimate**	**p-value**	**Estimate**	**p-value**
Intercept	876.36	0.000	652.56	0.000
Stream Order 2	−105.95	Not significant	75.35	Not significant
Stream Order 3	Not in data	–	333.39	0.002
Stream Order 4	Not in data	–	300.78	0.011
Land Use (Garden)	Not in data	–	12.16	Not significant
Land Use (Grassland)	Not in data	–	−167.64	Not significant
Land Use (Moorland)	Not in data	–	−129.10	Not significant
Land Use (Road)	−122.70	Not significant	−76.94	Not significant
Land Use (Woodland)	14.05	Not significant	−75.74	Not significant
Exposed (Yes)	−320.08	0.007	105.17	Not significant
R-squared	0.49		0.45	
Adjusted R-squared	0.36		0.28	
LTS Regression				
Intercept	995.68	0.000	828.88	0.000
Stream Order 2	−212.37	0.011	16.02	Not significant
Stream Order 3	Not in data	–	375.31	0.000
Stream Order 4	Not in data	–	325.60	0.003
Land Use (Garden)	Not in data	–	−153.20	Not significant
Land Use (Grassland)	Not in data	–	−378.35	0.012
Land Use (Moorland)	Not in data	–	−241.57	Not significant
Land Use (Road)	−242.02	Not significant	−278.08	Not significant
Land Use (Woodland)	−10.99	Not significant	−243.37	0.090
Exposed (Yes)	−278.84	0.003	95.92	Not significant
R-squared	0.70		0.59	
Adjusted R-squared	0.61		0.46	

effects were expected with the local TP SDs and MADs, as the same set of bandwidths were used. Global (non-spatial) TP means and medians were 702.1 and 701.0 mg kg^{-1}, respectively. Global (non-spatial) TP SDs and MADs were 241.4 and 295.6 mg kg^{-1}, respectively.

Local TP averages vary spatially along the stream network, where the nature of this variation depends on the average (mean or median) and distance metric (‘crow flies’ or stream) used (Fig. 5). The Monte Carlo tests (blue circles in Fig. 5) identify areas of unusually low and unusually high channel bank material TP concentrations. From Fig. 6, variability in TP itself also varies along the stream network, where the nature of this change in variation again depends on the variance measure (SD or MAD) and distance metric used. All maps in Fig. 6 are shown with the locations of an optimal design for a hypothetical future sampling campaign (also given as blue circles), for an example sample size of $n = 20$ (i.e., we have resources to visit 20 sites only for characterising channel bank material TP concentrations; (Collins et al., 2017)). Slightly different re-designs result, depending on how the local variance has been specified. Note that the re-designs are specific to the pre-defined set of regular points ($n = 110$), where these could easily have been set at a much finer scale.

3.7. Correlations for WEP

The results of the linear correlation (r) analysis to assess, statistically, the strength of WEP relationships to TP concentrations and to Fe + Al concentrations were undertaken on the $n = 40$ dataset associated with Catchment 3 (i.e., the widest spatially distributed sample set possible with minimal bias). The WEP concentrations were positively correlated to with TP concentrations ($r = 0.25$), but negatively correlated with Fe + Al concentrations ($r = -0.22$). Both correlations were significant at the 95% level.

4. Discussion

In their review of channel bank erosion rates and P concentrations, Fox et al. (2016) report that the values cited in the literature typically exceed 250 mg P kg^{-1}, while a more recent study has reported slightly lower values (171–304 mg P kg^{-1}) (Beck et al., 2018). Our values of channel bank TP, ranging between 129.6 and 1206.9 mg P kg^{-1} are comparable to previously reported data >250 mg P kg^{-1} (Fox et al., 2016). The range of values reported herein, however, is wider than some studies and more comparable to those reported by Kronvang et al. (2012) and Ishee et al. (2015). The levels of WEP in the channel bank samples ranged between 0.03 and 1.06 mg P kg^{-1} and comprised between 0.01 and 0.12% of TP. Few studies report water soluble or extractable P concentrations. Miller et al. (2014) found that WEP ranged between 1 and 25 mg P kg^{-1} and comprised, on average, between 0.03 and 5.1% of TP. Thoma et al. (2005) reported WEP concentrations ranging between 0 and 3.1 mg P kg^{-1} which comprised between 0 and 0.5% of TP. In the context of these studies, the WEP values reported here are lower. In common with other studies, the WEP concentrations of our channel bank samples were significantly related to TP and Fe + Al concentrations. The WEP concentrations were positively related to TP content and negatively related to Fe + Al content, both of which strongly bind with orthophosphate, limiting its availability. While the contribution of WEP from channel banks is likely to be negligible as a % of the total bank P contribution (e.g. Ross et al., 2019), ‘bioavailable’ sediment-associated P may be a potential source of P to the water column under anoxic conditions (Young and Ross, 2001).

Overall, this study shows that the increasing scale of the catchment sampled had limited effect on the variability of channel bank TP concentrations where only marginal evidence of different TP distributions was found between the full catchment (Catchment 3) and a nested catchment (Catchment 2). However, in contrast to the findings of Peacher et al. (2018), there was evidence that Stream Order did have a significant effect on TP values. This evidence was only at the full catchment scale where all four stream orders were present. On balance, the sampling strategy used within this study, that specifically focused on stream order coverage, was considered sufficient to capture the TP variability at all of the monitored scales and as such, stream order should be considered an important component of future channel bank TP sampling design so as to ensure the effects are captured. Although not implemented in this study, the optimal TP sample designs presented (and discussed below), could be constrained to ensure adequate representation of stream order, where sample numbers of $n = 10$ as used in this study, may need to be increased (depending on resources), accordingly.

While a significant relationship between grassland and woodland land use was found using a robust regression (and for only the dataset associated with Catchment 3), all other analyses indicated that riparian land use was not a significant factor in determining channel bank TP concentrations which is in common with the findings of Zaimes et al. (2008b) and Peacher et al. (2018), while both Miller et al. (2014) and Purvis et al. (2016) found that contrasting channel bank P concentrations can occur between similar streams in the same region even with similar channel bank characteristics, land use, and management. It is worth noting that the ‘Land Use’ classification was not a straightforward

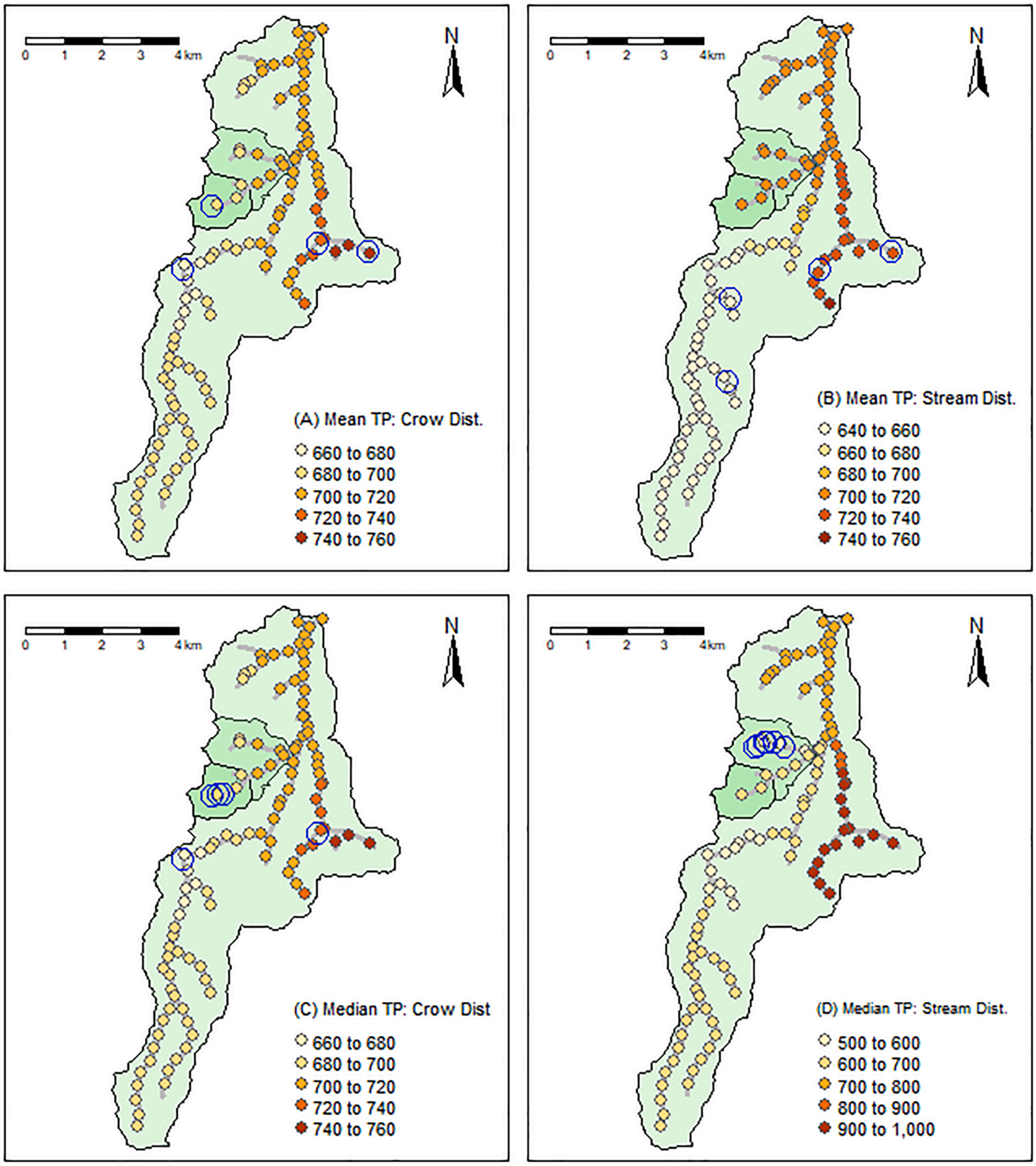

Fig. 5. Stream network local means (A and B) and local medians (C and D) found with different distance metrics - 'As the crow flies' (A and C) and 'Stream network' (B and D). Maps shown with Monte Carlo test results at the 95% level (blue circles at measured TP sites only).

assessment at sampling sites given the changing scale of the channels within the watershed. For example, small channels were often 'open' within a field with limited 'buffering' between them and the main adjacent land use. Larger channels were deep and frequently fenced off and had varying degrees of riparian buffer between the bank and the utilised agricultural land behind. When a tree lined bank ceased to be a 'woodland' and started to be a grassland or arable field was subjective. Similarly, at some sites, a grassland buffer was often left between the channel and arable land but given the variability of the geography of these features, it was not always clear whether the land use should be 'grassland' or 'arable' in these situations. A far clearer general distinction in land use can be made however, between the upland extensive grass and heathlands in the south of Catchment 3 and the more intensive agricultural land to the north. Critically, however, this land use gradient again showed no significant relationship to channel bank sample TP concentrations. To this end, and certainly within the lowland southern part of the catchments, the reason Land Use did not appear to affect channel bank P is that a large number of sites were distinct and separate from the land use above them and that the classification was largely falsely imposed upon the sample sites. Indeed, it might be argued that the main land use of the sampled channel banks was in fact 'channel bank'.

There is evidence to suggest that whether a channel bank is exposed/eroding affected the TP concentrations, at least within Catchment 2 of this study and therefore should be considered when sampling. This was probably an effect caused by TP concentrations changing with soil depth, (e.g. Haygarth et al., 1998; Miller et al., 2014). Within Catchment 2, this effect is probably strongly highlighted in samples termed 'Exposed' as a result of the changing morphology of the channel network with scale. In Catchment 1, all channels were Stream Order 1, generally shallow, and comprised mainly topsoil. In Catchment 2, however, the channels tended to be deeper (up to 5 m deep) and incised down into the clay rich subsoil layers. The significantly lower TP concentrations found in these Exposed samples (Fig. 4e) likely represents an increased contribution of subsoil from actively eroding taller channel banks. In other parts of the larger Catchment 3, this effect was not prevalent likely due to the river having incised its channel down into the bedrock, leaving less or no subsoil exposed. Given this, the contribution of TP

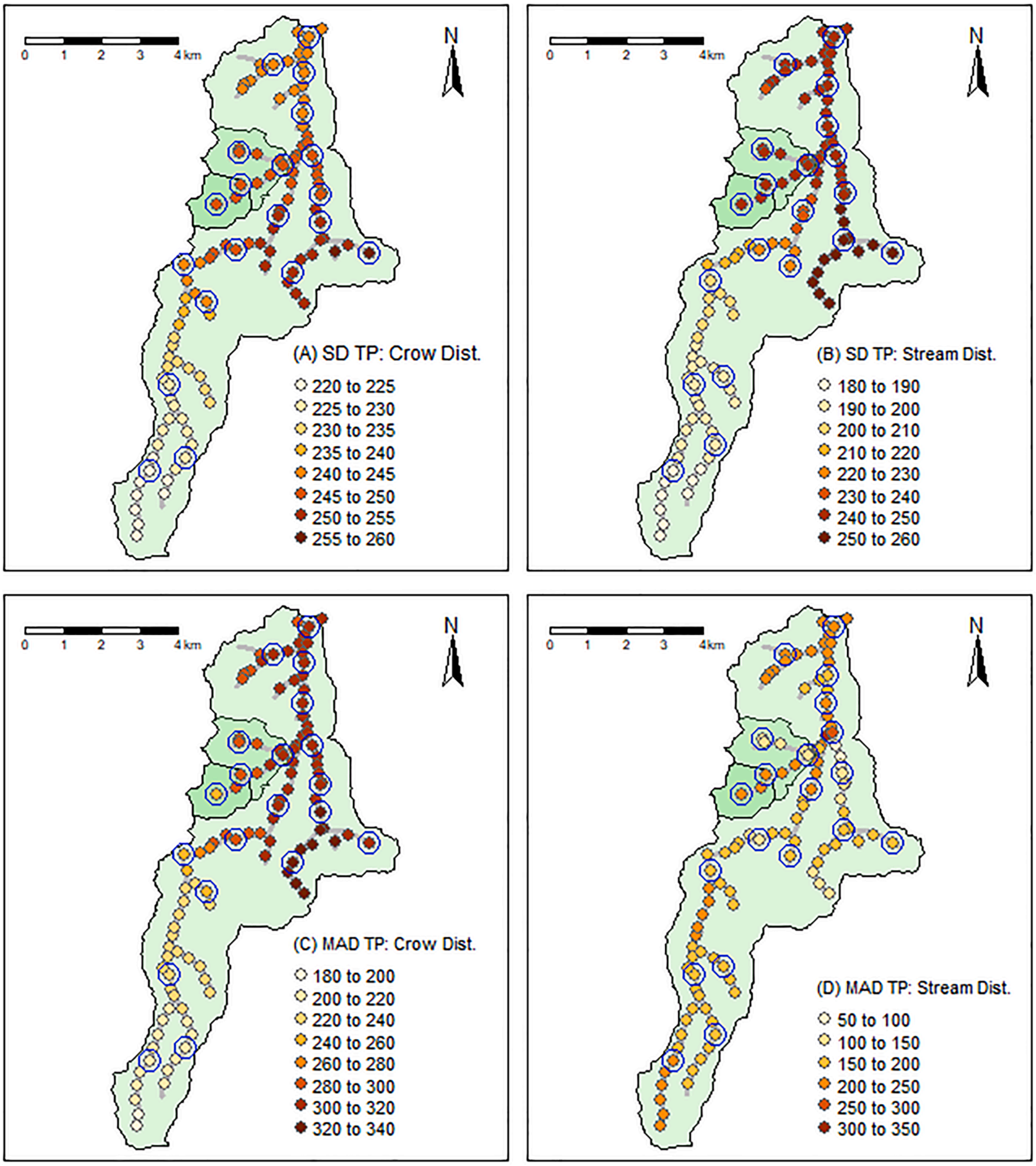

Fig. 6. Stream network local SDs (A and B) and local MADs (C and D) found with different distance metrics - 'As the crow flies' (A and C) and 'Stream network' (B and D). Maps shown with locations of an optimal design for a future sample campaign of size of $n = 20$ (blue circles at pre-defined set of regular points only).

from second order stream banks may be over-estimated given that when they are contributing sediment-associated P, the concentration may well be lower than that measured in banks that are *not* actively eroding. In the context of sampling channel banks for a source fingerprinting investigation, it is likely that exposed bank materials would be sampled and the TP content of those samples measured to provide an estimate of bank erosion derived contributions to the fluvial TP load (e.g., Walling et al., 2008).

The results of the spatial analysis demonstrated that the average and the variability of TP in local channel bank sediment vary spatially along the sampled stream network, where the nature of this variation depends on the statistic form (standard or robust) and distance metric ('crow flies' or stream) used. With respect to the most accurate spatial predictor, local TP means were more accurate than local TP medians, and the use of 'crow flies' rather than stream distances provides more accurate predictions; as indicated by their respective cross-validation scores, above. Thus, in a prediction sense, local TP means using Euclidean distances are the most accurate (Fig. 5A). However, as already stated, the main aims of this study were to examine changes in TP variability and here the local TP variances from which the TP sample re-designs are based are more pertinent. Given the local TP means using Euclidean distances are the most accurate, then it is reasonable to assume the local TP SDs using Euclidean distances provide the best representation of local TP variability, as both local measures share the same kernel bandwidth.

If local TP prediction accuracy was the only study goal, then the initial sampling strategy would be different in preparation for a *geostatistical* analysis with kriging (e.g. Chiles and Delfiner, 1999), rather than the *geographically weighted* analysis shown here. This is because, it is well known that kriging will be the more accurate predictor in such comparisons (e.g. Cressie, 1989). That said, as the study design was focused on stream order representation resulting in sample sites being spatially clustered (see Fig. 1), the design was not entirely ideal for any spatial analysis, geographically weighted (for local averages and variances), geostatistical or otherwise. For our spatial study, there was no ideal solution to the resultant bias (as described in Section 2.3.4), as bias would vary spatially, coupled with the tendency for a spatial analysis to suffer more than a non-spatial one through sparse information. This meant that any form of data sub-setting, as used in the non-spatial analyses, was not considered. Similarly, adopting a de-clustering approach where data in spatial clusters are assigned de-clustering weights to address the bias (e.g., Granger et al., 2017) was not considered, as it would also reduce information. In this respect, a pragmatic route was followed wherein the original n = 60 dataset was used, together with a caveat on the interpretation of the spatial results due to the described bias.

Given this caveat, the spatial analysis (Figs. 5 and 6) found the highest channel bank sample TP concentrations tended to be upstream in the middle south-east part of the stream network, while the lowest TP concentrations tended to be upstream in the bottom south-west part. Through Monte Carlo tests, areas were also identified with unusually high or unusually low TP concentrations (Fig. 5). Local TP variability appeared highest upstream to the middle south-east part of the stream network, except for local MADs using stream distances, which tended to indicate TP variability was highest downstream to the northern part of the network. Thus, interestingly, local TP means and SDs did appear to scale directly (as for many environmental processes), with the exception being local TP medians and MADs using stream distances. For example, high medians coincided with the low MADs upstream in the middle south-east part of the stream network when using stream distances. Crucially, the optimal designs for future sampling campaigns were found to be fairly robust to the choice of variance measure and the choice of distance metric, as only slight differences in designs were found (Fig. 6).

Although, it is usual in many catchment studies to test for outliers and associated non-normality for a given elemental composition (Collins et al., 2012, 2013), and then adopt a single path of presenting either a standard or a robust set of analyses only, this approach would have proved problematic in this study. This is because, many tests would need to be applied, given the different set of objectives for the statistical models used (e.g., an outlier identified for an ANOVA may not be outlying for a regression analysis), some of which would need to be spatial (e.g., Harris et al., 2014a). Further, testing for outliers / non-normality in a non-spatial manner may not suit the inherently spatial processes of this study. In this respect, both standard and robust analyses were presented together, where strong similarities in the results were taken as confirmation that outliers were not present or if present, were not influential, while strong dissimilarities were taken as confirmation of the opposite. This dual path approach was shown to be pertinent to the outcomes of both our non-spatial (e.g. the contrary regression results in Table 3) and spatial analyses.

5. Conclusions

Since channel bank erosion releases both sediment and associated P, this component of catchment systems should be included in source apportionment work using either empirical or modelling approaches. The TP concentrations measured in our channel bank samples are like those reported by the limited number of studies documenting such data. Variability in channel bank TP concentrations was primarily affected by Stream Order, and to a limited degree by how ‘exposed’ the sampled banks were, but not by Land Use or Catchment Scale. Sampling the larger of the three nested catchments alone would therefore likely be sufficient to capture TP variability provided stream order was included in the sampling strategy design. Local channel bank TP averages and TP variances varied spatially along the sampled stream network, where the most accurate spatial predictor of TP being local TP means using ‘crow flies’ distances while local TP variances provided inputs for optimal designs for future channel bank sampling.

Our work provides a novel demonstration of how commonly applied analyses of variance and regression complemented by models from a spatially-explicit geographically weighted framework can be used to investigate TP content in channel banks and help inform sampling protocols for subsequent studies. Such work is useful since the assembly of data on channel bank TP contents often derives from the sampling of banks for other purposes, rather than for explicitly and robustly representing P concentrations *per se*. This study has shown, in conjunction with the few other studies that examine channel bank P, that there are no fixed factors that can exclusively explain TP concentrations. Factors such as land use, or stream order may be significant in one catchment and not in another depending on the specific characteristics that define any given stream network. What our study has shown is that it would be wise to undertake a preliminary survey to determine where and which variables should be included and excluded from any such survey. Such preliminary work should highlight areas of the catchment where TP variability is highest and enable more strategic sampling to be undertaken. Fortunately, it would appear as though these preliminary surveys should capture all this information at large scales, or at least the largest scale undertaken within this study. Future work could also synthesize and extend the modelling framework of this study through the adaptation and implementation of a multiscale geographically weighted regression where different distance metrics can be assigned to different factors that drive variation in channel bank TP at different catchment scales (Lu et al., 2017).

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

Rothamsted Research receives strategic funding from UKRI-BBSRC (UK Research and Innovation – Biotechnology and Biological Sciences Research Council) and the work reported in this paper was undertaken as part of the institute strategic programme Soil to Nutrition project Work Package 3 (grant BBS/E/C/000I0330). The authors also acknowledge the landowners who permitted access to the channel network, the help of Mafalda Micale in sample collection and processing, and the assistance of Deborah Beaumont for information on the channel vegetation composition. The R software version R-3.6.3 (R Core Team, 2019) was used for the implementation of the statistical models, where the LTS regression was implemented using the robustbase R package (Maechler et al., 2020); the geographically weighted models and associated Monte Carlo tests were implemented using the GWmodel R package (Gollini et al., 2015); the network distances were found using the riverdist R package (Tyres, 2017); the regular sampling along the network was found using the spatialEco R package (Evans, 2020) and the location-allocation algorithm for the sample re-design was implemented using the tbart R package (Brunsdon, 2020).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.catena.2021.105223.

References

Beck, W., Isenhart, T., Moore, P., Schilling, K., Schultz, R., Tomer, M., 2018. Streambank alluvial unit contributions to suspended sediment and total phosphorus loads, Walnut Creek, Iowa, USA. Water 10, 22. https://doi.org/10.3390/w10020111.
Biddulph, M., Collins, A.L., Foster, I.D.L., Holmes, N., 2017. The scale problem in tackling diffuse water pollution from agriculture: insights from the Avon Demonstration Test Catchment programme in England. River Res. Appl. 33, 1527–1538. https://doi.org/10.1002/rra.3222.
Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geograph. Anal. 28, 281–298.
Brunsdon, C., Fotheringham, A.S., Charlton, M., 2002. Geographically weighted summary statistics — a framework for localised exploratory data analysis. Comput. Environ. Urban Syst. 26, 501–524.
Brunsdon, C., 2020. tbart: Teitz and Bart’s p-Median Algorithm R package version 1.0, http://tbart.r-forge.r-project.org/.

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.
Chiles, J., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley, New York.
Collins, A.L., Walling, D.E., Leeks, G.J.L., 1997. Sediment sources in the Upper Severn catchment: a fingerprinting approach. Hydrol. Earth Syst. Sci. 1, 509-521. https://doi.org/10.5194/hess-1-509-1997.
Collins, A.L., Zhang, Y., McChesney, D., Walling, D.E., Haley, S.M., Smith, P., 2012. Sediment source tracing in a lowland agricultural catchment in southern England using a modified procedure combining statistical analysis and numerical modelling. Sci. Total Environ. 414, 301-317. https://doi.org/10.1016/j.scitotenv.2011.10.062.
Collins, A.L., Zhang, Y.S., Hickinbotham, R., Bailey, G., Darlington, S., Grenfell, S.E., Evans, R., Blackwell, M., 2013. Contemporary fine-grained bed sediment sources across the River Wensum Demonstration Test Catchment, UK. Hydrol. Process. 27, 857–884. https://doi.org/10.1002/hyp.9654.
Collins, A.L., Pulley, S., Foster, I.D.L., Gellis, A., Porto, P., Horowitz, A.J., 2017. Sediment source fingerprinting as an aid to catchment management: a review of the current state of knowledge and a methodological decision-tree for end-users. J. Environ. Manag. 194, 86–108. https://doi.org/10.1016/j.jenvman.2016.09.075.
Conley, D.J., Paerl, H.W., Howarth, R.W., Boesch, D.F., Seitzinger, S.P., Havens, K.E., Lancelot, C., Likens, G.E., 2009. Ecology: Controlling eutrophication: nitrogen and phosphorus. Science 323, 1014–1015. https://doi.org/10.1126/science.1167755.
Cressie, N., 1989. The Many Faces of Spatial Prediction. In: Armstrong, M. (Ed.), Geostatistics. Proceedings of the Third International Geostatistics Congress. Springer Netherlands, Avignon, France, pp. 163-176.
Dunn, O.J., 1964. Multiple comparisons using rank sums. Technometrics 6, 241–252. https://doi.org/10.2307/1266041.
EEC, 2000. Directive 2000/60/EC of the European parliment and the council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Commiunties. L327, 1-72.
Evans, B.M., Sheeder, S.A., Lehning, D.W., 2003. A spatial technique for estimating streambank erosion based on watershed characteristics. J. Spatial Hydrol. 3, 1–13.
Evans, J.S., 2020. spatialEco. R package version 1.3-1, https://github.com/jeffreyevans/spatialEco.
Finn, R.N., 2007. The physiology and toxicology of salmonid eggs and larvae in relation to water quality criteria. Aquat. Toxicol. 81, 337–354. https://doi.org/10.1016/j.aquatox.2006.12.021.
Fox, G.A., Purvis, R.A., Penn, C.J., 2016. Streambanks: a net source of sediment and phosphorus to streams and rivers. J. Environ. Manag. 181, 602–614. https://doi.org/10.1016/j.jenvman.2016.06.071.
Gallacher, K., Miller, C., Scott, E.M., Willows, R., Pope, L., Douglass, J., 2017. Flow-directed PCA for monitoring networks. Environmetrics 28, 9. https://doi.org/10.1002/env.2434.
Gollini, I., Lu, B.B., Charlton, M., Brunsdon, C., Harris, P., 2015. GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models. J. Stat. Softw. 63, 1–50.
Granger, S.J., Harris, P., Peukert, S., Guo, R., Tamburini, F., Blackwell, M.S.A., Howden, N.J.K., McGrath, S., 2017. Phosphate stable oxygen isotope variability within a temperate agricultural soil. Geoderma 285, 64–75. https://doi.org/10.1016/j.geoderma.2016.09.020.
Haddadchi, A., Ryder, D.S., Evrard, O., Olley, J., 2013. Sediment fingerprinting in fluvial systems: review of tracers, sediment sources and mixing models. Int. J. Sediment Res. 28, 560–578. https://doi.org/10.1016/s1001-6279(14)60013-5.
Harris, P., Brunsdon, C., 2010. Exploring spatial variation and spatial relationships in a freshwater acidification critical load data set for Great Britain using geographically weighted summary statistics. Comput. Geosci. 36, 54–70. https://doi.org/10.1016/j.cageo.2009.04.012.
Harris, P., Brunsdon, C., Charlton, M., Juggins, S., Clarke, A., 2014a. Multivariate spatial outlier detection using robust geographically weighted methods. Math. Geosci. 46, 1–31. https://doi.org/10.1007/s11004-013-9491-0.
Harris, P., Clarke, A., Juggins, S., Brunsdon, C., Charlton, M., 2014b. Geographically weighted methods and their use in network re-designs for environmental monitoring. Stoch. Environ. Res. Risk Assess. 28, 1869–1887. https://doi.org/10.1007/s00477-014-0851-1.
Haygarth, P.M., Hepworth, L., Jarvis, S.C., 1998. Forms of phosphorus transfer in hydrological pathways from soil under grazed grassland. Eur. J. Soil Sci. 49, 65–72.
Haygarth, P.M., ApSimon, H., Betson, M., Harris, D., Hodgkinson, R., Withers, P.J.A., 2009. Mitigating diffuse phosphorus transfer from agriculture according to cost and efficiency. J. Environ. Qual. 38, 2012–2022. https://doi.org/10.2134/jeq2008.0102.
Heathwaite, A.L., Dils, R.M., 2000. Characterising phosphorus loss in surface and subsurface hydrological pathways. Sci. Total Environ. 251, 523–538.
Heathwaite, A.L., Dils, R.M., Liu, S., Carvalho, L., Brazier, R.E., Pope, L., Hughes, M., Phillips, G., May, L., 2005. A tiered risk-based approach for predicting diffuse and point source phosphorus losses in agricultural areas. Sci. Total Environ. 344, 225–239. https://doi.org/10.1016/j.scitotenv.2005.02.034.
Imeson, A.C., Vis, M., Duysings, J.J.H.M., 1984. Surface and subsurface sources of suspended solids in forested drainage basins in the Keuper region of Luxembourg. In: Burt, T.P., Walling, D.E. (Eds.), Catchment experiments in fluvial geomorphology. GeoBooks, Norwich, U.K., pp. 219–234
Ishee, E.R., Ross, D.S., Garvey, K.M., Bourgault, R.R., Ford, C.R., 2015. Phosphorus characterization and contribution from eroding streambank soils of Vermont's Lake Champlain basin. J. Environ. Qual. 44, 1745–1753. https://doi.org/10.2134/jeq2015.02.0108.
Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., Brook, J.R., 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. Atmos. Environ. 39, 2399–2409. https://doi.org/10.1016/j.atmosenv.2004.06.049.
Kronvang, B., Laubel, A., Grant, R., 1997. Suspended sediment and particulate phosphorus transport and delivery pathways in an arable catchment, Gelbæk Stream. Denmark. Hydrol. Process. 11, 627–642.
Kronvang, B., Vagstad, N., Behrendt, H., Bogestrand, J., Larsen, S.E., 2007. Phosphorus losses at the catchment scale within Europe: an overview. Soil Use Manag. 23, 104–116. https://doi.org/10.1111/j.1475-2743.2007.00113.x.
Kronvang, B., Audet, J., Baattrup-Pedersen, A., Jensen, H.S., Larsen, S.E., 2012. Phosphorus load to surface water from bank erosion in a Danish lowland river basin. J. Environ. Qual. 41, 304–313. https://doi.org/10.2134/jeq2010.0434.
Laubel, A., Kronvang, B., Hald, A.B., Jensen, C., 2003. Hydromorphological and biological factors influencing sediment and phosphorus loss via bank erosion in small lowland rural streams in Denmark. Hydrol. Process. 17, 3443–3463. https://doi.org/10.1002/hyp.1302.
Lloyd, C.E.M., Johnes, P.J., Freer, J.E., Carswell, A.M., Jones, J.I., Stirling, M.W., Hodgkinson, R.A., Richmond, C., Collins, A.L., 2019. Determining the sources of nutrient flux to water in headwater catchments: examining the speciation balance to inform the targeting of mitigation measures. Sci. Total Environ. 648, 1179–1200. https://doi.org/10.1016/j.scitotenv.2018.08.190.
Lu, B., Charlton, M., Harris, P., Fotheringham, A.S., 2014. Geographically weighted regression with a non- Euclidean distance metric: a case study using hedonic house price data. Int. J. Geogr. Inf. Sci. 28, 660–681. https://doi.org/10.1080/13658816.2013.865739.
Lu, B., Charlton, M., Brunsdon, C., Harris, P., 2017. Geographically weighted regression with parameter-specific distance metrics. Int. J. Geogr. Inf. Sci. 31, 982–998. https://doi.org/10.1080/13658816.2016.1263731.
Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Conceicao, E.L., Anna di Palma, M., 2020. robustbase: Basic Robust Statistics. R package version 0.93-6, http://robustbase.r-forge.r-project.org/.
McDowell, R.W., Dils, R.M., Collins, A.L., Flahive, K.A., Sharpley, A.N., Quinn, J., 2016. A review of the policies and implementation of practices to decrease water quality impairment by phosphorus in New Zealand, the UK, and the US. Nutr. Cycl. Agroecosyst. 104, 289-305. https://doi.org/10.1007/s10705-015-9727-0.
Meybeck, M., Helmer, R., 1989. The quality of rivers - from pristine stage to global pollution. Glob. Planet. Change 75, 283–309. https://doi.org/10.1016/0921-8181(89)90007-6.
Miller, R.B., Fox, G.A., Penn, C.J., Wilson, S., Parnell, A., Purvis, R.A., Criswell, K., 2014. Estimating sediment and phosphorus loads from streambanks with and without riparian protection. Agric. Ecosyst. Environ. 189, 70–81. https://doi.org/10.1016/j.agee.2014.03.016.
Minella, J.P.G., Merten, G.H., Walling, D.E., Reichert, J.M., 2009. Changing sediment yield as an indicator of improved soil management practices in southern Brazil. Catena 79, 228–236. https://doi.org/10.1016/j.catena.2009.02.020.
Morse, G.K., Laster, J.N., Perry, R., 1993. The economic and environmental impact of phosphorus removal from wastewater in the European Community. Selper Publications, London.
Murphy, J., Riley, J.P., 1962. A modified single solution method for determination of phosphate in natural waters. Anal. Chim. Acta 26, 31–36.
Nair, V.D., 2014. Soil phosphorus saturation ratio for risk assessment in land use systems. Front. Environ. Sci. 2 https://doi.org/10.3389/fenvs.2014.00006.
Owens, P.N., Walling, D.E., Leeks, G.J.L., 2000. Tracing fluvial suspended sediment sources in the catchment of the River Tweed, Scotland, using composite fingerprints and a numerical mixing model. In: Foster, I.D.L. (Ed.), Tracers in Geomorphology. John Wiley & Sons Ltd, Chichester, pp. 291–308.
Peacher, R.D., Lerch, R.N., Schultz, R.C., Willett, C.D., Isenhart, T.M., 2018. Factors controlling streambank erosion and phosphorus loss in claypan watersheds. J. Soil Water Conserv. 73, 189–199. https://doi.org/10.2489/jswc.73.2.189.
Purvis, R.A., Fox, G.A., Penn, C.J., Storm, D.E., Parnell, A., 2016. Estimating streambank phosphorus loads at the watershed scale with uncertainty analysis approach. J. Hydrol. Eng. 21, 12. https://doi.org/10.1061/(asce)he.1943-5584.0001402.
Rode, M., op de Hipt, F., Collins, A.L., Zhang, Y.S., Theuring, P., Schkade, U.K., Diekkruger, B., 2018. Subsurface sources contribute substantially to fine-grained suspended sediment transported in a tropical West African watershed in Burkina Faso. Land Degrad. Dev. 29, 4092-4105. https://doi.org/10.1002/ldr.3165.
Roseboom, D.P., 1987. Case studies of stream and river restoration. in: Urbana-Champaign, U.o.I.a. (Ed.), Management of the Illinois River System: The 1990's and Beyond : Proceedings : a Governor's Conference on April 1-3, 1987 at Peoria, Illinois for Citizens, Organizations, Industry, and Government Representatives and Resources Management Professionals. Water Resources Center, University of Illinois, pp. 184-194.
Rosgen, D.L., 1994. A classification of natural rivers. Catena 22, 169–199. https://doi.org/10.1016/0341-8162(94)90001-9.
Ross, D.S., Wemple, B.C., Willson, L.J., Balling, C.M., Underwood, K.L., Hamshaw, S.D., 2019. Impact of an extreme storm event on river corridor bank erosion and phosphorus mobilization in a mountainous watershed in the northeastern United States. J. Geophys. Res.-Biogeosci. 124, 18–32. https://doi.org/10.1029/2018jg004497.
Rousseeuw, P.J., Leroy, A.M., 1987. Robust regression and outlier detection. John Wiley & Sons.
Rousseeuw, P.J., Croux, C., 1993. Alternatives to the median absolute deviation. J. Am. Stat. Assoc. 88, 1273–1283. https://doi.org/10.2307/2291267.
Royer, T.V., David, M.B., Gentry, L.E., 2006. Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: implications for reducing

nutrient loading to the Mississippi River. Environ. Sci. Technol. 40, 4126–4131. https://doi.org/10.1021/es052573n.
Sekely, A.C., Mulla, D.J., Bauer, D.W., 2002. Streambank slumping and its contribution to the phosphorus and suspended sediment loads of the Blue Earth River. Minnesota. J. Soil Water Conserv. 57, 243–250.
Slattery, M.C., Walden, J., Burt, T.P., 2000. Use of mineral magnetic measurements to fingerprint suspended sediment sources: Results from a linear mixing model. John Wiley & Sons Ltd, Chichester.
Stone, M., Mudroch, A., 1989. The effect of particle-size, chemistry and mineralogy of river sediments on phosphate adsorption. Environ. Technol. Lett. 10, 501–510. https://doi.org/10.1080/09593338909384766.
Stone, M., English, M.C., 1993. Geochemical composition, phosphorus speciation and mass-transport of fine-grained sediment in 2 Lake Erie tributaries. Hydrobiologia 253, 17–29. https://doi.org/10.1007/bf00050719.
Taylor, A.W., Kunishi, H.M., 1971. Phosphate equilibria on stream sediment and soil in a watershed draining an agricultural region. J. Agric. Food Chem. 19, 827–831. https://doi.org/10.1021/jf60177a061.
Teitz, M.B., Bart, P., 1968. Heuristic methods for estimating generalized vertex median of a weighted graph. Oper. Res. 16, 955–961. https://doi.org/10.1287/opre.16.5.955.
Thoma, D.P., Gupta, S.C., Bauer, M.E., Kirchoff, C.E., 2005. Airborne laser scanning for riverbank erosion assessment. Remote Sens. Environ. 95, 493–501. https://doi.org/10.1016/j.rse.2005.01.012.
Tukey, J.W., 1949. Comparing individual means in the analysis of variance. Biometrics 5, 99–114. https://doi.org/10.2307/3001913.
Tyres, M., 2017. riverdist: River network distance computation and applications R package version 0.15-0, http://riverdist.r-forge.r-project.org/.
Vargha, A., Delaney, H.D., 1998. The Kruskal-Wallis test and stochastic homogeneity. J. Educ. Behav. Stat. 23, 170-192.
Ver Hoef, J.M., Peterson, E., Theobald, D., 2006. Spatial statistical models that use flow and stream distance. Environ. Ecol. Stat. 13, 449–464. https://doi.org/10.1007/s10651-006-0022-8.
Ver Hoef, J.M., Peterson, E.E., 2010. A moving average approach for spatial statistical models of stream networks. J. Am. Stat. Assoc. 105, 6–18. https://doi.org/10.1198/jasa.2009.ap08248.
Walling, D.E., Owens, P.N., Leeks, G.J.L., 1999. Fingerprinting suspended sediment sources in the catchment of the River Ouse, Yorkshire, UK. Hydrol. Process. 13, 955–975.
Walling, D.E., Collins, A.L., Stroud, R.W., 2008. Tracing suspended sediment and particulate phosphorus sources in catchments. J. Hydrol. 350, 274–289. https://doi.org/10.1016/j.jhydrol.2007.10.047.
Young, E.O., Ross, D.S., 2001. Phosphate release from seasonally flooded soils: a laboratory microcosm study. J. Environ. Qual. 30, 91–101. https://doi.org/10.2134/jeq2001.30191x.
Zaimes, G.N., Schultz, R.C., Isenhart, T.M., 2008a. Streambank soil and phosphorus losses under different riparian land-uses in Iowa. J. Am. Water Resour. Assoc. 44, 935–947. https://doi.org/10.1111/j.1752-1688.2008.00210.x.
Zaimes, G.N., Schultz, R.C., Isenhart, T.M., 2008b. Total phosphorus concentrations and compaction in riparian areas under different riparian land-uses of Iowa. Agric. Ecosyst. Environ. 127, 22–30. https://doi.org/10.1016/j.agee.2008.02.008.
Zhang, Y., Collins, A.L., Murdoch, N., Lee, D., Naden, P.S., 2014. Cross sector contributions to river pollution in England and Wales: updating waterbody scale information to support policy delivery for the Water Framework Directive. Environ. Sci. Policy 42, 16–32. https://doi.org/10.1016/j.envsci.2014.04.010.