Phosphorus buffering in streams by benthic sediments

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy

at
Lincoln University
by
Zachary P. Simpson

Lincoln University
2021

Abstract of a thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

Phosphorus buffering in streams by benthic sediments

by

Zachary P. Simpson

The loss of phosphorus (P) to aquatic ecosystems accelerates eutrophication – a problem felt worldwide. Central to any effort to monitor and mitigate the effect of P in stream is knowing how inputs of P, whether point or diffuse, map to P transport downstream. However, the stream itself possesses several mechanisms to attenuate P inputs thus blurring the connection between P inputs and P availability in-stream. For example, various stream biota and geochemical processes may remove P from or even release P to the water column. In particular, benthic stream sediments have the capacity to sorb P to their surfaces which may later desorb back into solution. This P sorption means benthic sediments can behave much like a buffer for P: a transient store of P which may offset changes to P concentrations in the stream. The thesis of this work is that the benthic sediment-P buffer is a predominant control on P availability at baseflow in streams. In its five studies, I investigate sediment-P sorption in detail but also examine multiple alternative pathways for stream P retention. Special attention is given to the sediment equilibrium phosphate concentration at net zero sorption (EPC_0), which is the dissolved reactive P (DRP) concentration towards which sediments buffer DRP concentrations in the solution (i.e., sediment porewater and, via hyporheic exchange, the water column) through sorption.

A systematic review and meta-analysis of the EPC_0 in streams at baseflow – covering 45 studies and 466 stream sites across the globe – found wide variability in the disparity between in-stream DRP concentrations and sediment EPC_0 (termed as a potential for P exchange). This contrasts with previous views that P in sediments and streamwater is typically in an ‘equilibrium’. Further, this potential for P exchange was moderated by sediment and stream characteristics, including sorption affinity, pH, and sediment exchangeable P concentrations. For example, fine benthic sediments are often highly sorptive but may also restrict hydrological exchange between the water column and the hyporheic zone, leading to wider disparities. Methodological factors were also influential (e.g., choice of solution, sediment pre-treatment, equilibration time), indicating a need for research on unified EPC_0 methodology.

The second study established that drying sediments prior to analyses (either air- or freeze-drying) biases sediment P fractionation measurements and inflates the variance of EPC_0. Such drying techniques may lyse microbial biomass P, alter organic P availability, and age metal oxides responsible for sorption, thus

complicating the natural sediment P chemistry. Instead, analyzing stream sediments fresh (wet) is recommended.

The third study surveyed a variety of stream waters and sediments from catchments draining three distinct lithologies (alluvium, sedimentary, and volcanic basic) to assess the likelihood of various geochemical controls on stream DRP concentrations. Geochemical equilibria in the water-column indicated no significant potential for the (co-)precipitation of minerals that could sequester P (e.g., calcite). However, the sediments stored large amounts of P in labile and redox-sensitive forms: indeed, this labile P correlated with stream DRP concentrations but only for streams with likely sufficient hyporheic exchange. Catchment geology and redox cycling in stream influence sediment reactivity for P and so are a major source of between-stream variability in DRP concentrations.

The fourth study focused on a confounding factor when interpreting EPC_0: is P sorption or microbial P cycling responsible for sediment P buffering? Unlike some previous work, this study found a minimal role for sediment microbes to alter sediment EPC_0 values even with replete labile carbon and nitrogen sources available. Further, sediments sterilized via γ-irradiation did increase in EPC_0, but this increase was attributed to lysis of the microbial biomass – an overlooked P stock in streams. The study highlights that the sediment P buffer, while largely abiotic in nature, may subsidize microbial P demand in sediment biofilms, thus influencing stream ecological function.

The last study examined P uptake at the stream reach scale. Considering two contrasting but predominant controls on stream P uptake – periphyton P demand and sediment P sorption – a natural way to separate the two processes was to measure P uptake under light and dark conditions. Stream gross primary productivity (driven by periphyton) was high as expected for this open-canopy stream. However, this did not translate to an increase in stream P uptake when compared to dark conditions. Sediments were highly sorptive and their relatively low EPC_0 suggested a potential for P removal throughout the experiment. Thus the sediment P buffer was likely most responsible for the measured stream P uptake although different stream conditions (e.g., greater nitrogen availability) could increase periphyton's relevance and should be studied further.

In summary, the benthic sediment P buffer can contribute to the regulation of P availability in many streams. Sediments may attenuate P inputs, thus dampening DRP variation at baseflow and prolonging the legacy of past P inputs in the catchment. Stream hydrology (e.g., hyporheic exchange), geochemistry (e.g. pH), and biota (e.g., sediment microbial P demand) are among the chief external factors that may moderate or interact with the sediment P buffer and deserve further study. Predicting P availability in streams remains a major challenge, but understanding the sediment P buffer will greatly improve our ability to prevent stream eutrophication.

Keywords: phosphate, sorption, rivers, water quality, eutrophication, lotic, periphyton, hyporheic zone, ecological stoichiometry, legacy P.

Acknowledgements

I am deeply grateful for the opportunity to pursue a PhD. I first thank my PhD supervisors: Professors Rich McDowell and Leo Condron.

Rich gave me an incredible opportunity for this study as well as an immense amount of freedom in defining the project. Rich's keen mentorship kept me inquisitive yet focused, helping me to craft a thesis that wasn't an anthology of dead ends. I've taken inspiration from Rich as a scientist: broad-minded, unafraid to tackle new topics, thinks across all scales, deeply motivated about solving real problems, pragmatic. I will continue to sharpen my thinking on Rich whether he's present in person or on paper. I've also taken inspiration from Rich as a human: the most authentic modern-day Stoic I know, incredibly aware, yet simultaneously a whimsical goofball. Rich, I can't thank you enough for the past few years.

Leo has been the best co-supervisor possible. I naively assumed he was a Kiwi before arriving and was pleasantly surprised when I met the Glaswegian Gabber in person. Leo is undoubtedly hilarious but is more importantly a great scientist and a wonderful mentor. It is abundantly clear that Leo gives his students full respect and I believe that makes all the difference. Leo, thanks for being my sounding board and an inexhaustible pool of encouragement.

Thanks to the Our Land and Water National Science Challenge, whose funding made this PhD project possible. Special thanks also to Matt Iremonger at Willesden Farms for providing great access to Kaituna River. The review and meta-analysis presented in Chapter 2 was made possible by all the researchers whose studies were included; in addition, helpful discussions, insight, and critical reviews provided by Marshall McDaniel, Jonathan Abell, and Helen Jarvie greatly sharpened the study.

Special thanks to the global R community – all statistical analyses, data figures, and maps in this thesis were made possible by the R programming language which is supported by the legions of scientists working towards making science open, transparent, and (in my opinion at least) more enjoyable.

The staff at Lincoln are thanked for all their technical help and for making the soils department what it is, particularly Roger Cresswell, Leanne Hassall, and Lynne Clucas.

I thank the rest of the fellow Lincoln post-grad & friends population for making Lincoln a better place: Yuan Li, David 'Krafty Kraut' Rex, Marion des Roseaux, Thomas Néron, Andrea Leptin, Renato Ricciardi (the aforementioned were revolving denizens of an unforgettable flat), Gustavo Boitt, Phuong Nguyen, Carolina Lizzaralde, Camilla Gardiner, Dharshika Welikala & Keshana, Tihana Vujinović, Camille Rousset, Luciano Nunes Leite, Zicheng Yi, Ying Zhao, Jonathan Nuñez & Adriana Medina, Daniel Hendrie, Carmen Medina Carmona & Michael Brown, and Shyam Provost. Thanks also to David and Diane Whitehead for their hospitality and kindness.

I especially thank Professor Brian Haggard for encouraging me to pursue a PhD to begin with and for his continued mentorship as well as Gustavo Boitt for taking the time to get me started in the lab, the discussions about all things phosphorus, the (mis-)adventures, and for encouraging me when I most needed it.

Thanks to my family for their patience and support across hemispheres: Mom, Dad, Mallory, Luke, Grandma, and Grandpa.

Lastly, I want to especially thank Andrea: a bottomless well of kindness and support, my best friend, and the love of my life. Dankeschön für alles.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

1.1 The problem with phosphorus in streams

Phosphorus (P) is a key element of life. In cells across the kingdoms of life, P is uniquely vital for much of the biomolecular machinery (ribosomes), genetic material (DNA, RNA), energy currency (ADP/ATP), and more (Sterner and Elser 2002; Kamerlin et al. 2013). As a consequence, P is also vital for agriculture. Worryingly, the world's agricultural sector is expected to support some 9-10 billion people by 2050 despite increasingly scarce (or economically unfeasible) sources of phosphate (Smil 2000). This underscores two major goals for P sustainability: maximize P use efficiency and minimize P losses. My focus is on the latter.

The loss of P from land is itself a crucial problem for two reasons. First, as mentioned, loss of added P means P not used for food production. Second, P loss damages aquatic ecosystems by accelerating primary production, i.e., eutrophication. Without being exhaustive in listing all the problems, eutrophication degrades biodiversity (Evans-White et al. 2009; Ardón et al. 2021), diminishes the aesthetic value of our waters, costs untold billions of dollars annually (Dodds et al. 2009), and threatens access to potable water for millions of people (Wurtsbaugh et al. 2019). Preventing P loss to waters is therefore paramount. Any solution towards tightening the P economy as well as restoring impaired water bodies will hinge on knowing when, where, and how P moves through catchments and their stream networks.

Phosphorus-limited or nitrogen and P co-limited eutrophication is widespread and globally affects the waters that support billions of people (McDowell et al. 2020b). Aquatic ecosystems are particularly sensitive to dissolved reactive P (DRP; the filterable proportion of P that is operationally defined as 'reactive' and correlates strongly with bioavailability; Biggs 2000; Dodds 2007). Hence, monitoring programs heavily invest in measuring DRP across streams and through time. The hope is that such data may inform of our impact on P loads or concentrations and perhaps the efficacy of our management. The challenge in interpreting such monitoring data, though, is that any grab sample collected at a fixed point in time and space will contain DRP that has (1) derived from multiple natural and anthropogenic inputs (Schlesinger and Bernhardt 2013; Dupas et al. 2018a; McDowell et al. 2019a); (2) travelled along a wide distribution of flow paths (Burt and Pinay 2005; McDonnell 2013); (3) travelled for a wide distribution of travel times (Rinaldo et al. 2015; Sprenger et al. 2019); and (4) cycled through numerous biotic and abiotic compartments along the way (House 2003; Haggard and Sharpley 2007; Withers and Jarvie 2008). Of these components, management usually centers on the first (anthropogenic P inputs), but all components blend together and thus obscure the effects of changing P inputs. More simply, our observations of P in streams and rivers are functions of catchment attributes (components 1 through 3)

and of in-stream processing (the 4th component of this challenge). The motivation of the present thesis is that by elucidating in-stream processing of P, the remaining obstacles with disentangling P pollution may be more surmountable.

It is my thesis that, as a key component of in-stream P processing, benthic sediments contribute to the regulation of the transport of P through streams especially at baseflow. More accurately, *benthic sediments act as a buffer for DRP*. Below I will overview P transport through streams at baseflow conditions and highlight the function of sediments in P transport; this will be brief because a much larger exposition will be given in Chapter 2. Then I will state the objectives of this thesis and its organization.

1.2 Overview of P cycling within streams and the relevance of the sediment P buffer

Once in the stream, P may undergo many transformations (Haggard and Sharpley 2007; Withers and Jarvie 2008; Records et al. 2016). Following the seminal work of Newbold (1983) and Mulholland (1983, 1994) and the plethora of stream tracer studies following the paradigm set out by Stream Solute Workshop (1990), P transport in streams is often viewed as a 'spiral'. That is, DRP may be taken up by biota (e.g., algae) which may then feed invertebrates or grazing fish and so forth before being released again to the water column either as waste or through microbial mineralization (Meyer 1994); at the same time, DRP may undergo abiotic transformations that removes it from the water column, such as mineral (co-)precipitation or complexation with particles, and may later be released back into the water-column. (Please refer to Figure 1 of Stream Solute Workshop (1990) for a helpful conceptual diagram.) The challenge with most stream P cycling work in this nutrient spiraling paradigm is that measurements of P uptake and release at the stream scale, reviewed in meta-analyses by Ensign and Doyle (2006) and by Hall et al. (2013), are unspecific about the responsible mechanisms, whether biotic and abiotic.

Here I will just focus on some prominent processes with emphasis on what may occur at baseflow and make the case for why sediment P sorption (what I term the sediment P buffer) is my focus – this will not be exhaustive (Stream Solute Workshop 1990; Reddy et al. 1999; Haggard and Sharpley 2007; Withers and Jarvie 2008). An important challenge to bear in mind is that these processes often occur simultaneously, meaning that parameterizing the responsible or even dominant P cycling mechanisms in streams remains a major research need. However, accurate process attribution is key to our understanding of stream P transport and, therefore, to our predictions of how DRP in the water column responds to changing P inputs upstream.

1.2.1 Biotic P pathways

I will highlight two particular biotic stocks as dominant in stream P cycling: benthic periphyton and sediment microbes. Sometimes both are discussed together as 'biofilms' since, indeed, the two are well-connected (Battin et al. 2016). Periphyton (also known as aufwuchs) are a complex mixture of microorganisms which form a biofilm on stream substrate. This biofilm can grow into a variety of mats

or other structures where autotrophic microbes are the dominant identifiable feature (Biggs 1996), but within the mat often exists a complex microbiome of many heterotrophs and autotrophs. Periphyton, which are themselves part of the eutrophication problem, can influence the amount of DRP available in streams through their own P demand (Mulholland et al. 1994; Dodds 2003).

Microbes entrained within the sediment are similarly an important P sink but differ considerably from periphyton. While perhaps a part of the biofilm continuum (Battin et al. 2016), microbes underneath the surface substrate are in the dark waters of the hyporheic zone of the stream (where surface and ground water interact and exchange with the stream; Boano et al. 2014). Hence, these sediment microbes are usually dominated by heterotrophs, particularly bacteria, which also demand P if their carbon and nitrogen demands are met (Hill et al. 2012; Sinsabaugh et al. 2012). Sediment microbes may also indirectly interact with P cycling in streams such as by encouraging the reductive dissolution of iron oxides, which are important to P sorption (Lovley 1991; Smolders et al. 2017).

1.2.2 Abiotic P pathways

Much of the P in stream DRP is often some form of orthophosphate ($H_2PO_4^-$ and/or HPO_4^{2-} in most streams) – though DRP should not be equated with phosphate – and phosphate is immensely reactive in the natural environment. I will again highlight just two notable mechanisms: mineral precipitation and sorption.

With other cations, phosphate can form many mineral phases which may then lower the activity of phosphate in solution until equilibrium is reached (Lindsay et al. 1989; Stumm and Morgan 1996). This is particularly well-studied in soils, where the equilibria with minerals such as hydroxylapatite (a calcareous phosphate mineral) can alter solution P (Lindsay et al. 1989; McDowell and Sharpley 2003a), but certain stream environments may also bring about P exchange between mineral and solution phases, particularly within benthic sediments (Machesky et al. 2010; Tye et al. 2016). Calcite deserves special mention since it is common in the waters of many calcareous catchments, interacts dynamically with stream pH (and hence photosynthetic activity; Stumm and Morgan 1996; Stets et al. 2017), and can co-precipitate with phosphate (Plant and House 2002; Sø et al. 2011). For example, Corman et al. (2016) shaded a calcium-rich stream reach, reducing its calcite deposition rate (by 57%) due to lack of periphyton photosynthesis, leading to significantly less DRP uptake compared to an unshaded control. Analogously, dynamic redox fronts in some streams, particularly in the hyporheic zone, could encourage the formation of various ferric phosphate co-precipitants (Senn et al. 2015). Clearly, the geochemical setting of the stream will determine the mineral equilibria relevant to P cycling.

Phosphate may also bind to mineral surfaces through adsorption. Being one of the most sorptive (oxy)anions (Sigg and Stumm 1981; Sposito 2004), phosphate readily complexes with particle surfaces in both soils and sediments. Sediments rich in hydrous metal oxides (particularly Fe and Al) can rapidly adsorb and sequester large amounts of phosphate (van der Grift et al. 2014; McDowell 2015). Other

reactive surfaces common to sediments can act as sorption sites as well (clay minerals, some carbonates), depending on the sediment's geochemical makeup (Machesky et al. 2010; Rawlins 2011). Indeed, sorption is likely the principal mechanism behind the sediment uptake of P (Froelich 1988). The near ubiquity of benthic sediments in streams, especially in agricultural catchments, and their large storage capacity for P make sediments one of the largest, transient stores of P in streams.

Yet, benthic sediments are often overlooked among investigations of stream DRP dynamics. Even in reviews/meta-analyses of nutrient cycling in streams, sediment P sorption receives scant attention (Ensign and Doyle 2006; Hall et al. 2013). This negligence may later be built into our models of stream P transport, therefore hindering progress on predicting the fate and transport of P.

1.2.3 The sediment P buffer and its difficulty

I focus on sediments not only because they can adsorb and store such large quantities of P, but because sediments may also *desorb* P back to the water column. This dynamic of P sorption in transient sediment storage is what this thesis focuses on: the sediment P buffer. The sediment P buffer implies that streams may remove P under some conditions only to later release it back to the water-column, thus tending to maintain DRP concentrations at baseflow analogous to how a pH buffer resists changes to H^+ activity (Taylor and Kunishi 1971; Froelich 1988). Since the sediment P buffer may attenuate DRP concentrations at baseflow so effectively, it can make P pollution mitigation measures appear inconsequential (Stutter et al. 2010; Meals et al. 2010). Hence, the sediment P buffer is clearly one of the more difficult aspects of P cycling in streams and remains a major research gap (Hamilton 2012).

Particularly, most research to date on the stream sediment P buffer has focused on a metric termed the equilibrium phosphate concentration at net zero sorption, or the EPC_0. The EPC_0 is effectively the 'level' of the sediment P buffer (sticking with the pH buffer analogy, a buffer solution may be made up to have a level of pH 8 and so forces more acidic or more alkaline solutions towards this level). That is, sediments tend to buffer solution DRP towards the sediment EPC_0. However, like for the pH buffer, inputs eventually can overpower the buffer and move the system to a new equilibrium. While the sediment P buffer is not defined by the EPC_0, the metric is closely related. A full literature review and meta-analysis of the EPC_0 is given in Chapter 2 along with a broader review of the sediment P buffer in general.

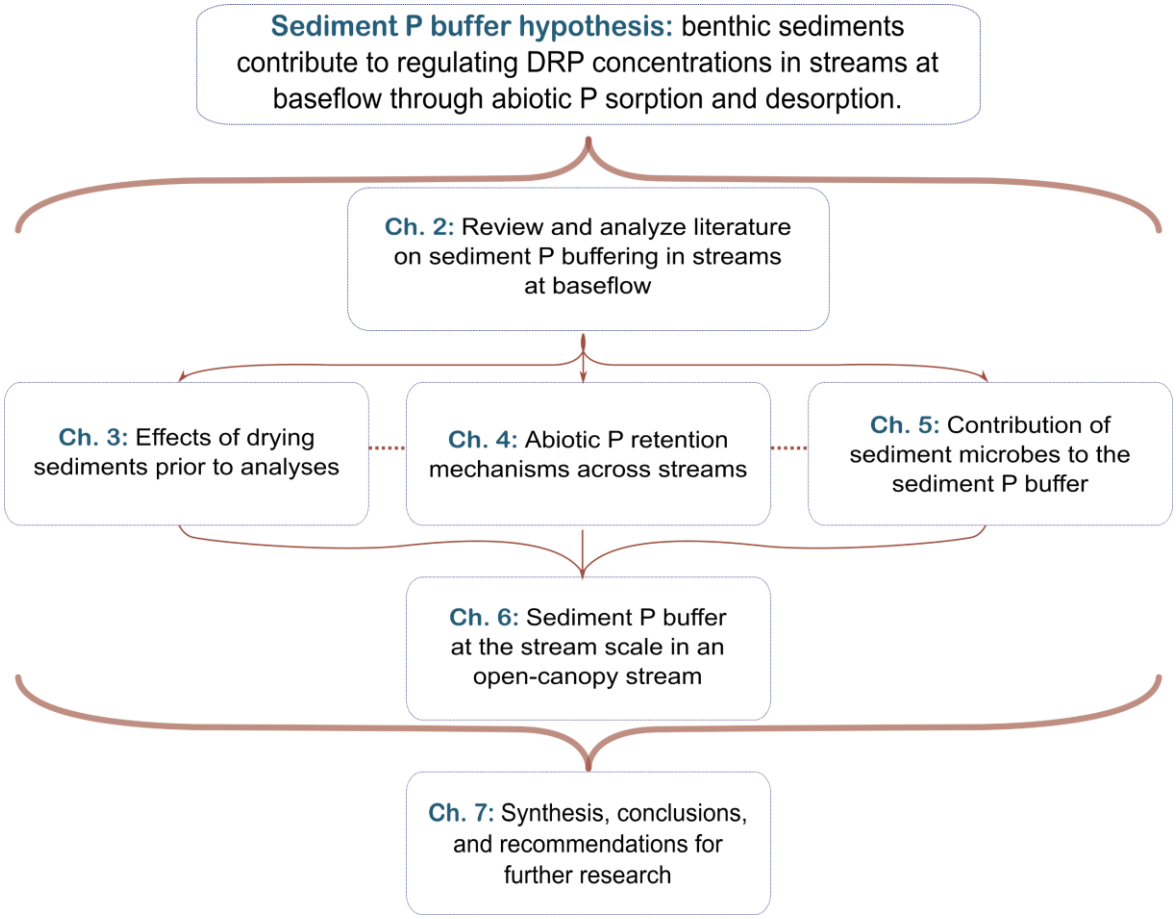

Figure 1.1 General structure of the thesis.

1.3 Thesis objectives and structure

The central hypothesis of this work was that, through abiotic P sorption, benthic stream sediments contribute substantially to the regulation of DRP concentrations at baseflow, much like a buffer. This hypothesis informed the objectives and the structure of the thesis (Figure 1.1). My objectives were concerned with the reasons for and implications of this sediment P buffer and aimed to:

I. Quantitatively synthesize the existing literature on the sediment P buffer to arrive at comprehensive conclusions on how sediments buffer P, which covariates are important, experimental considerations for future work, and test whether, in general, sediment P and water-column DRP are typically at an 'equilibrium'.
II. Test whether sediment drying, as a pretreatment method prior to analyses, affects sediment P fractions and EPC_0 relative to analyzing sediments 'fresh' or wet.
III. Survey stream waters and benthic sediments across a variety of stream environments to explore relationships between potential abiotic P retention mechanisms and the geochemistry of sediments and the water-column.
IV. Test the apparent contribution of sediment microbes to the sediment P buffer.
V. Distinguish two major pathways for P retention at the whole-stream scale: periphyton P uptake and sediment P sorption.

Objectives I through V correspond to chapters 2 through 6. Chapter 2 provides an in-depth review of the sediment P buffer but also is itself an analysis. This chapter highlights further research gaps, some of which are addressed by chapters 3 through 6, and provides context for the remainder of the thesis. Chapter 3 is a practical methodological study and highlights an important facet of sediment P analyses. Benefitting from the wide variety in the streams of Canterbury, New Zealand, chapter 4 studies the abiotic (i.e., geochemical) mechanisms responsible for P retention in streams with special attention given to sediments. Chapter 5 is still focused on sediments but challenges the central thesis by questioning whether sediment microbes are responsible for part of the sediment P buffer. Drawing from Chapters 2 through 5, Chapter 6 is an intensive in-stream study that attempts to isolate the contribution of sediments to P retention at the stream-reach scale. Chapter 7 synthesizes the findings of the thesis, states conclusions, and provides consideration for future research. Further material supplementary to the research chapters are available as appendices A through E.

Chapter 2
Sediment phosphorus buffering in streams at baseflow: A meta-analysis

2.1 Abstract

Phosphorus (P) pollution of surface waters remains a challenge for protecting and improving water quality. Central to the challenge is understanding what regulates P concentrations in streams. This quantitative review synthesizes the literature on a major control of P concentrations in streams at baseflow – the sediment P buffer – to better understand streamwater-sediment P interactions. We conducted a global meta-analysis of sediment equilibrium phosphate concentrations at net zero sorption (EPC_0), which is the dissolved reactive P (DRP) concentration towards which sediments buffer solution DRP. Our analysis of 45 studies and >900 paired observations of DRP and EPC_0 showed that sediments often have potential to remove or release P to the streamwater (83% of observations), meaning that 'equilibrium' between sediment and streamwater is rare. This potential for P exchange is moderated by sediment and stream characteristics including: sorption affinity, stream pH, exchangeable P concentration, and particle sizes. The potential for sediments to modify streamwater DRP concentrations is often not realized owing to other factors, e.g., hydrologic interactions. Sediment surface chemistry, hyporheic exchange, and biota can also influence the potential exchange of P between sediments and the streamwater. Methodological choices significantly influenced EPC_0 determination and thus the estimated potential for P exchange: we therefore discuss how to measure and report EPC_0 to best suit research objectives and aid in inter-study comparison. Our results enhance understanding of the sediment P buffer and inform how EPC_0 can be effectively applied to improve management of aquatic P pollution and eutrophication.

2.2 Introduction

> *"The base material of subsoils and stream banks has a very large capacity to adsorb phosphate and **acts as a strong buffer** to reduce the phosphate level of the stream."*
>
> *Taylor and Kunishi, 1971 (emphasis added)*

Phosphorus (P) pollution of surface waters is a global problem, where excess P may fuel eutrophication (Elser et al. 2007; Dodds and Smith 2016). At baseflow, where stream eutrophication is most pronounced (Biggs 1996, 2000; Dodds 2006), benthic sediments can buffer (retain or release) dissolved reactive P (DRP, primarily orthophosphate; Haggard et al. 1999; Small et al. 2016; Griffiths and Johnson 2018), the main form of P bioavailable to primary producers including periphyton (Biggs 2000; Wetzel 2001; Muscarella et al. 2014). Hence, to predict and improve water quality, it is important to understand how the sediments buffer the supply of P from inputs upstream and to stream biota.

The most common method used to quantify the sediment P buffer is the equilibrium phosphate concentration at net zero sorption (EPC_0, mg P L^{-1}; see also Appendix A for the general method). The EPC_0 is the DRP concentration in solution at which sediment will neither adsorb nor desorb P from the surrounding solution under laboratory conditions (Taylor and Kunishi 1971; Froelich 1988). It is also the DRP concentration where sediments display the greatest buffering capacity (Froelich, 1988). Consequently, DRP concentrations above and below the sediment EPC_0 suggest a potential for P adsorption and desorption, respectively. Originally developed for soils (White and Beckett 1964), the EPC_0 concept was first applied to stream sediments by Taylor and Kunishi in 1971; since then, numerous studies have applied the EPC_0 concept (Figure 2.1).

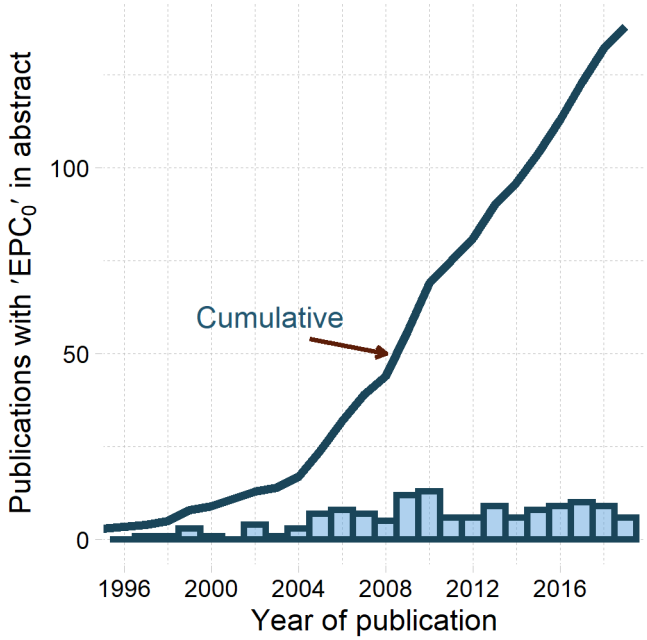

Figure 2.1 Frequency of papers published per year with the term 'EPC_0' in the abstract through the year 2019. The Scopus search string (for this figure only) was 'ABS(epc0*)'; notably, this search overlooks many studies that only report benthic stream sediment EPC_0 within the text.

At baseflow, EPC_0 describes how benthic sediments potentially buffer DRP concentrations. The difference between sediment EPC_0 and in-stream DRP concentrations indicates the direction and potential magnitude of sediment P sorption. Across many streams, sediment EPC_0 correlates well with in-stream DRP concentrations (Jarvie et al. 2005; Haggard et al. 2007; McDaniel et al. 2009; Machesky et al. 2010; McDowell 2015). The correlation between DRP and EPC_0 is likely a consequence of how sediment-streamwater interactions regulate P supply. Point-source inputs of P provide a clear

example of this, whereby sediment EPC_0 observed along a stream network often increases downstream concurrently with the increasing DRP concentrations (Ekka et al. 2006; Jarvie et al. 2006b; Roberts and Cooper 2018; Wilcock et al. 2020). Likewise, sediments also buffer non-point P inputs from intensive land uses, which then gets reflected in sediment EPC_0 (Haggard et al. 2007; Stutter and Lumsdon 2008; Agudelo et al. 2011; McDowell 2015). Sediments themselves can be the vector for losses of legacy P (Sharpley et al. 2013), connecting P bound to soils on land to the storage of P on streambeds (McDowell et al. 2003; Agudelo et al. 2011; Emelko et al. 2016). No matter the P source, these relationships with EPC_0 suggest that, across catchments, benthic sediments attenuate P losses by regulating DRP concentrations. The ability of benthic sediments to buffer high DRP concentrations, however, diminishes as the EPC_0 increases. P previously sorbed to such sediments may desorb back to the water-column when conditions change (Haggard et al. 2005; Haggard 2010; Stutter et al. 2010). Processes like the sediment P buffer serve to dampen P loads to downstream ecosystems but cannot necessarily prevent the long-term downstream flux of P.

This DRP buffering by sediments is apparent across a range of scales, from smaller streams such as agricultural ditches (Kröger and Moore 2011; Ezzati et al. 2020) to large streams and rivers (Jarvie et al., 2005; McDaniel et al., 2009; McDowell et al., 2015). For this reason, EPC_0 is used in some catchment- or stream-scale P transport models, where EPC_0 is a parameter in models estimating benthic sediment P fluxes (van der Perk 1997; House and Warwick 1999; House and Denison 2002; White et al. 2014; Jackson-Blake et al. 2015). The sediment P buffer, however, is not static. For example, EPC_0 can change following storm events or changes in sediment sources due to new sediments depositing on the streambed (Emelko et al. 2016; McDowell et al. 2019b). Excessive DRP loading could overwhelm the capacity of the sediment P buffer, weakening the P uptake potential in the stream (Meyer 1979; Marti et al. 2004; Weigelhofer et al. 2018b). In a way, sediment EPC_0 integrates past biogeochemical changes in the stream network including the history of sediment exposure to DRP (Jarvie et al. 2005; Hamilton 2012), which may influence current, in-stream P cycling.

Studies on benthic sediment EPC_0 cover a wide variety of environments, pollution sources, and scales, yet there has been little work to integrate these findings on how the sediment P buffer mechanism mediates downstream DRP supply. Part of the challenge may be in the interpretation of EPC_0 and the lack of standardized EPC_0 methodology. For example, the solution ionic strength, Ca concentration, and pH can alter the measured EPC_0 (Meyer 1979; Barrow 1983a; Klotz 1988; Wang et al. 2006; Lucci et al. 2010; Bhadha et al. 2012; Huang et al. 2016). Recent research has not always accounted for these factors. Another issue arises in interpreting EPC_0 measurements. Most researchers compare sediment EPC_0 to water-column DRP, inferring whether the sediment is a source or sink for water-column DRP. However, sediment sorption sites may be hydrologically disconnected from the water column in some cases (Weigelhofer et al. 2018a), and so 'active' uptake or release should instead be termed as a *potential* for sediment P adsorption or desorption (Hoffman et al. 2009; Stutter et al. 2010). We contend that kinetic limitations on the sediment P buffer are often involved (Taylor and Kunishi, 1971; Froelich,

1988) and that insightful interpretations of in-stream P cycling requires additional hydrological and biogeochemical information (Covino 2017; Marcé et al. 2018).

To aid in our critical review of EPC_0, we conducted a meta-analysis following a systematic literature review. Such a review is challenging for two primary reasons. First, meta-analyses typically require consistent, summary effect sizes (and their errors) designed for a given hypothesis (e.g., response ratios and correlation coefficients; Rosenberg et al. 2013) across the studies under review (Koricheva and Gurevitch 2014; Nakagawa et al. 2017; Gurevitch et al. 2018). Such a standardized effect size does not exist for EPC_0, which is usually measured at specific points in time and space alongside other variables in a more observational study design. Hence, in a meta-analysis of EPC_0, we need to use the primary data (i.e., individual observations; Mengersen et al. 2013a). Second, a myriad of factors or moderator variables (Nakagawa et al., 2017; Gurevitch et al., 2018) influence EPC_0 and its relation to water-column DRP, but these are inconsistently measured and reported throughout the literature. To model the effects of these moderators on the disparity between EPC_0 and DRP, traditional statistical methods rely on complete datasets. Consequently, where information on moderator variables is incomplete, data would be omitted and our inferences biased (Little and Rubin 2002; Nakagawa and Freckleton 2011).

To test hypotheses regarding the sediment P buffer, we address these challenges by standardizing the primary data – paired observations of EPC_0 and DRP – with a new metric termed the phosphate exchange potential (PEP). The PEP reflects the difference in magnitudes of EPC_0 and DRP, whereby greater differences from zero indicate a potential for P exchange (PEP is discussed in more detail below). Further, we overcome problems with missing data by applying multiple imputation (Rubin 1996; Van Buuren 2018), thus making the most of all information available.

In this review, we critically examine the most common applications of EPC_0 by testing three hypotheses (H_1-H_3):

1. EPC_0 and DRP will tend towards the same value at baseflow, i.e., global PEP values will distribute closely about zero if potential for P exchange is realized during sufficiently long, stable conditions. In addition to testing whether the statistical expectation of the PEP distribution equals zero, we compare the dataset to a threshold of 'equilibrium' used by Jarvie et al. (2005).

2. EPC_0, but not necessarily PEP, changes when P inputs to the stream change. To test this hypothesis, we analyze measurements taken concurrently on the same stream but above and below a point source (e.g., a wastewater treatment plant), which should yield similar PEP, as we assume the processes controlling any difference between DRP and EPC_0 to be equivalent at each point.

3. The variation in PEP can be explained by catchment, stream, and sediment physiochemical characteristics and the methodological variables controlled by the investigators. To examine

this, we build a statistical model to predict PEP based on key variables reported in the systematic review.

Less commonly reported, but integral for understanding the sediment P buffer, we explored those few studies that measured seasonal variability in EPC_0 within individual streams and those that assessed the contribution of sediment microbiota to apparent sediment P uptake. Too few data were available in this review for traditional meta-analytical techniques so we address these hypotheses more directly:

4. Variability in PEP within a stream over time is driven by variation in environmental factors that vary seasonally (e.g., temperature, light, stream discharge, and sediment inputs). We assess this hypothesis by collating and examining data from streams where three or more measurements over time were available.

5. Sediment microbial processes (i.e., P assimilation and mineralization) contribute to stream sediment P uptake and release. We assess this hypothesis by collating and examining sediment P sorption assays conducted using fresh and sterilized/microbially-inhibited sediments.

Through our analyses, we refine our conceptualization of the stream sediment P buffer, clarify its critical role in stream P transport, and recommend directions for future research on the topic.

2.3 Materials and Methods

2.3.1 Systematic Literature Review

We conducted a systematic literature search for papers on benthic sediment EPC_0 in lotic, freshwater environments. To make our study rigorous and transparent, we adhered to the PRISMA guidelines (Preferred Reporting Items for Systematic reviews and Meta-analyses; Liberati et al. 2009) as closely as possible. Not all PRISMA items were addressed, including review protocol registration, due to the exploratory nature of the source papers and of our review.

Full details of our search method and criteria for inclusion are in Appendix A. Briefly, we searched the Web of Science and Scopus databases in July 2018 for papers containing concurrent data for benthic sediment EPC_0 and in-stream DRP. We only considered peer-reviewed publications written in English. From our initial result of 1807 papers (excluding repeats), we screened papers by title, abstract, and then by full-text according to our criteria for inclusion (Côté et al. 2013), yielding 48 papers. Three of these papers did not yield data since either the paper did not include data or the corresponding author was unreachable. Hence, our dataset synthesizes 45 papers (Table A.1). Due to the systematic nature of our search (Nakagawa et al., 2017; Gurevitch et al., 2018), future studies can replicate and supplement our work with more source papers (e.g., non-English publications and future publications).

2.3.2 Data Curation and Handling for Meta-Analysis

Our meta-analysis differs from many others in that we considered the primary data (here, individual measurements of EPC_0 and DRP for a given point in space and time) rather than a reported effect size and associated error, as is common in other meta-analytic literature (Mengersen et al., 2013; Rosenberg et al., 2013). We focused on the primary data since studies with EPC_0 generally do not summarize results into an overall effect (i.e., corresponding to a specific hypothesis). While our use of primary data increases the complexity of our study design, it is also a strength as we can compare study-specific methods and scrutinize relationships among various moderator variables (Mengersen et al., 2013). Therefore, we extracted individual observations of concurrent EPC_0 and DRP, in addition to related stream and sediment physicochemistry, site characteristics, and experimental methods (Table A.2). We collated data from target papers by: recording data given in tables, extracting data from figures via the Engauge tool (Mitchell et al. 2020; http://markummitchell.github.io/engauge-digitizer/), and by contacting corresponding authors. Full details of the data collection and handling are in Appendix A.

To facilitate imputations of missing values and improve parameter estimates in subsequent modelling (see below), we $\log_{10}$-transformed variables having long-tailed distributions: sediment total P, OC, fines concentration, linear sorption coefficient (K_d), exchangeable P, catchment area, and equilibration time.

2.3.3 Effect Metric for Sediment Phosphorus Buffering: Phosphate Exchange Potential (PEP)

Previous work quantified the disparity between sediment EPC_0 and DRP with various metrics. Jarvie et al. (2005) used EPC_{sat} (%), defined as $100\times(EPC_0\text{-DRP})/EPC_0$; Gainswin et al. (2006) used a phosphorus transfer index (PTI), defined as $(DRP/EPC_0 -1)$; and Son et al. (2015) used the difference between EPC_0 and DRP alongside EPC_{sat}. Metrics involving a ratio of un-transformed variables, such as EPC_{sat} and PTI can have poor statistical qualities (primarily skewness; Rosenberg et al., 2013). The arithmetic difference between EPC_0 and DRP is straightforward but is difficult to assess when comparing data from low-P streams (e.g., differences of ± 0.01 mg P L^{-1}) to that of polluted streams (e.g., differences of ± 1 mg P L^{-1}) – i.e., the scale needs to be accounted for to make all observations comparable. We therefore defined a metric, termed the *phosphate exchange potential* (PEP), which is similar to a log response-ratio common in meta-analyses:

$$\mathrm{PEP} = \log_{10}\left(\frac{\mathrm{EPC_0}}{\mathrm{DRP}}\right) = \log_{10}(\mathrm{EPC_0}) - \log_{10}(\mathrm{DRP})$$

The PEP reflects the difference in magnitudes of EPC_0 and DRP, or the potential for sediments to exchange P with the surrounding solution. PEP maintains the desired qualities of indicating direction (i.e., negative values indicate a potential for P removal by sediments while positive values indicate potential for sediment P release) while also accounting for differences in scale through the log-

transformation. Importantly, the PEP metric is a *potential*: other processes such as the variable kinetics of P sorption or the hydrologic exchange of waters may influence whether this potential is realized.

All PEP data were used to test our first hypothesis (H_1) that PEP would distribute about zero and our third hypothesis (H_3) that sediment, stream, and catchment characteristics and methodological variables moderate PEP (Table A.2). Studies with multiple sites on a stream with and without point-sources were used to test the second hypothesis (H_2; EPC_0 changes with changes in P inputs but PEP would not vary). For this test (see below), we selected observations immediately above and below the point-source, as reported.

In our secondary analyses, we examined stream sites with $n \geq 3$ observations over time, including multiple sites on the same stream (H_4). We also synthesized reported tests on the contribution of sediment biota to EPC_0 or sediment P sorption (H_5); these studies generally inhibited or sterilized sediment microbes and made comparisons in P sorption metrics with a fresh sediment.

2.3.4 Statistical tests for H_1 and H_2

For our first hypothesis (H_1; that the distribution of PEP tended towards zero, indicating an average zero potential for sediment sorbed P to alter DRP in the water column), we used the Wilcoxon signed-rank test on the full dataset. Our effect metric, PEP, aside from reducing skewness relative to metrics like EPC_{sat}, also improved the symmetry of the distribution, making the test for its location more robust (Hollander et al. 2013a). The alternative hypothesis was that the location of PEP was not zero, for which we estimated the shift in location and its 95% confidence interval.

We used a subset of the data meeting the requirements for our second hypothesis (H_2) that PEP varies along a stream due to a specific P input (i.e., a point-source). Additionally, we tested for the change in EPC_0 for these sediments, which required a log-transformation to improve its distributional symmetry. With these paired measurements, we used the Wilcoxon signed-rank test to test for a change in either EPC_0 or PEP downstream from a point-source.

2.3.5 Statistical Modeling of Moderators for PEP (H_3)

4.3.5.1 Missing Data and Multiple Imputation

For our third hypothesis (H_3, that PEP would depend on sediment, stream, and catchment characteristics as well as methodological variables), we analyzed many common variables reported alongside DRP and EPC_0 (Table A.2). We sought to use a type of meta-regression to explore possible moderators of PEP (Mengersen et al. 2013b; Gurevitch et al. 2018) but many of the moderator variables of interest were poorly reported throughout our systematic review (Figure A.1). Methods like regression rely on complete data; dropping incomplete rows of data to estimate a statistical model is untenable as, even for a modest subset of variables, the majority of the data would be unused and parameter estimates would be biased (Little 1992; Little and Rubin 2002; Nakagawa and Freckleton 2011). For example,

considering only stream pH and benthic sediment fines concentration as moderator variables, only 35% of the data would be useable; for the subset of variables we ultimately used in our model (see below), 0% of the data had complete cases.

Instead, we employed missing data techniques to best represent the information in the systematic review and develop our statistical model while incorporating all observed data. Specifically, we used multiple imputation (Rubin 1987, 1996; Little and Rubin 2002), a technique used in previous meta-analyses (Nakagawa and Freckleton 2011; Crane-Droesch et al. 2013; Jolani et al. 2015). Full details behind our multiple imputation procedure, via the MICE package (Van Buuren and Groothuis-Oudshoorn 2011; Van Buuren 2018), are in Appendix A. Briefly, we generated multiple plausible imputations for missing data to generate multiple (plausible) completed datasets. Imputations are based on individual predictive models that leverage information in the rest of the dataset to make reasonable imputations (summarized in Table A.2). For example, predictions for a missing sediment fines concentration can be better constrained by accounting for sediment organic C concentration (and vice versa) since the two variables tend to covary in streams (Findlay 1995; Tank et al. 2010). The process is repeated iteratively across all variables with missing observations before converging to one completed dataset. Each completed dataset (here, *m*=500 datasets) is then fitted by a statistical model. According to Rubin's rules (Rubin, 1987; Van Buuren, 2018), these *m* model fits were pooled to one model from which we base our inferences. Below we describe the statistical model used and the pooling procedure.

4.3.5.2 Generalized Additive Models (GAMs) for PEP

For each of the 500 completed datasets generated via multiple imputation, we fitted a model for PEP as a function of the available moderator variables. Based on our literature review, we knew several variables would have nonlinear relationships (e.g., between EPC_0 and sorption capacity as K_d; Haggard et al. 2007; Zhang and Huang 2007) and transformations would be of minor help. Therefore, we used generalized additive models (GAM; Hastie et al. 2009) via the 'mgcv' package (Wood 2011) as a flexible approach. Our final model used nonparametric smooth terms and parametric linear terms. GAMs were fitted by restricted maximum likelihood estimation and had the general form of:

$$PEP = \alpha + \tau_s + X\beta + \sum_{j=1}^{p} f_j(Z) + \varepsilon$$

where, α is an intercept term, τ_s is the random effect for study *s*, X is the vector of moderators that are modelled linearly with parameters β, Z is the vector of moderators that are modelled with p nonparametric smooth functions f, and ε is the residual error modelled here as Gaussian ($\sim N(0, \text{scale}^2)$).

Following best practice for meta-analyses, we included random effects at the study level (τ_s) to account for between-study heterogeneity in our models (Nakagawa et al., 2017; Gurevitch et al., 2018); random effects are implemented in the GAMs as a smooth term as described by Wood (2013a). For the parametric terms in X, we used moderator variables that were categorical or that were continuous but

had clustered values, including: solution ionic strength, $\log_{10}$ equilibrium time, sediment pretreatment, and upstream point-source influence. For the nonparametric smooth terms, which were fitted with cubic-regression splines, we used the following continuous variables: stream pH, $\log_{10}$ fines concentration, $\log_{10}$ exchangeable P, and $\log_{10} K_d$.

While all potential moderator variables informed our analysis through the multiple imputation procedure, only some were used in the GAM. We included as many relevant variables as possible in the GAM but culled several due to either (1) having poor information (indicated by missing information indices and unstable imputation distributions; see Appendix A and Figure A.5) or (2) indirectly influencing PEP but being collinear with other moderators (e.g., agricultural land use was dropped as it was collinear with sediment fines concentration). Multiple model formulations and diagnostics (residual plots, missing information statistics, GAM fit summary statistics) informed our final model selection, which we consider as the best compromise between the lack of information in the data and lack of explanatory power in moderator variables.

4.3.5.3 Pooling of Multiply-imputed GAMs

Each individual GAM was then pooled into one final model to complete the multiple imputation approach. This process is detailed in Appendix A and is similar to that described in Crane-Droesch et al. (2013).

Pooling provides the final model, its parameter estimates, and missing data statistics (Little and Rubin, 2002; Van Buuren, 2018), which are necessary in interpreting the model. In addition to parameter estimates, standard errors, and *p*-values, we also report: $\bar{U}$ (the within-imputation variance, i.e., the estimated sampling variance for the parameter estimates across the imputations); B (the between-imputation variance, i.e., the variance in parameter estimates due to missing data as reflected in the variable imputations); T_{var} as the total variance, estimated by $T_{var} = \bar{U} + ((m+1)/m) \times B$, and whose square root gives the parameter standard error; relative increase in variance (RIV), a measure of the proportional increase in total variance due to unobserved values; and fraction of missing information (FMI), which is the proportion of the total variance attributable to unobserved data (i.e., B).

Certain statistics could not be properly pooled, but we provide approximations (see Appendix A) which are denoted with *.

2.3.6 Data and code

All statistical analyses were performed in R (ver. 3.6.1; R Core Team 2020). The literature search results (and processing), collated dataset extracted from the systematic review, and the associated R code for analyses are available on Figshare (https://doi.org/10.6084/m9.figshare.12570059).

2.4 Results

2.4.1 Description of the data

The review resulted in 45 primary studies meeting the criteria, totaling 942 observations of paired EPC_0 and DRP concentrations across 9 countries (Figure 2.2). While many observations were concentrated in a few countries (Austria, New Zealand, United Kingdom, and United States), the environments and catchment types varied widely, providing a firm basis to infer about the disparity in DRP-EPC_0 data in temperate regions. However, we refrain from inferring about EPC_0 or the DRP-EPC_0 disparity in tropical or polar regions.

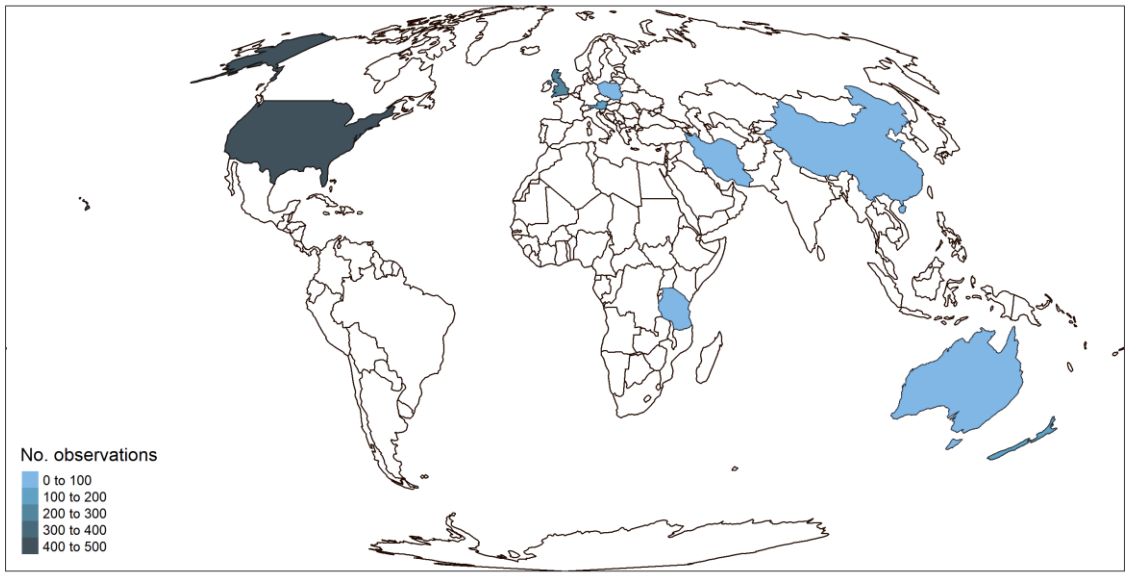

Figure 2.2 Location by jurisdiction for the source studies in the review (45 total). Total observations of paired DRP and EPC_0 by country are shown via color gradient.

Data collated from the review are summarized in Table A.2 and the missingness characteristics of the data are illustrated in Figure A.1. Catchment sizes ranged from small streams (0.24 km^2) to large rivers (20,582 km^2). Most catchments studied were primarily agricultural (median observed land use of 63%), with forest being the next common land use (30%). For observed cases (76% of the data), a third of the streams had at least one point-source upstream.

For stream chemistry, few variables were available or consistently reported alongside EPC_0 (e.g., stream pH was observed only 55% of the time). Reflecting the scope of the primary studies, sediment variables were more frequently reported but notably suffered from missingness depending on the specific research aims of the study. Sediment pH was poorly reported (11.6%), while ~50% of the data had corresponding measurements of some form of organic C (as either organic C or organic matter via loss on ignition, combined via Ball's equation (Ball 1964); range from 0 to 112 g kg^{-1}), fines concentration (0 to 983 g kg^{-1}), and total P (31 to 14,000 mg P kg^{-1}). Other sediment-P variables aside from EPC_0 were difficult to include due to changing methodologies (e.g., P fractions), but could be aggregated into two categories:

labile sediment P pools and P sorption affinity. Labile sediment P pools consisted of bioavailable P via the Fe-strip method (Sharpley 1993; 12% of the data) or an exchangeable P pool (either water- or salt-extractable P; 27% of the data). P sorption affinity was reported either by the Bache-Williams index (1.2% of data), a linear sorption isotherm slope (K_d; 24%), or by the Langmuir isotherm affinity parameter (k_{Lang}; 4.6%).

Methodologies for EPC_0 varied across all studies. Most sediments were analyzed fresh (92%), although a few were analyzed frozen/freeze-dried (1.4%) or air-dried (6.3%). Most studies also applied some form of solution correction to account for the background matrix (i.e., to mimic ambient streamwater chemistry; use of D.I. H_2O applied to 3.5% of data). Solution ionic strength averaged 0.008 ± 0.007 M (observed for 72% of data). Equilibration time for the batch experiments ranged from 1 h to 7 d, with a median of 18 h.

2.4.2 EPC_0, DRP, and their disparity: phosphate exchange potential (PEP)

As previously observed across several individual studies, DRP and EPC_0 were well correlated (Figure 2.3). Both variables covered similar ranges (<1 µg P L^{-1} to ~7 mg P L^{-1}) and were approximately log-normally distributed. After taking the difference between the two log_{10}-transformed variables, our effect metric – PEP – tended towards a normal distribution but varied widely (Figure 2.4). A QQ-plot of PEP is given in Figure A.6, illustrating the heavy tails in the distribution (slightly greater variation than expected for a normal distribution). Using a paired rank test, we estimated the location of PEP's distribution to be slightly negative (-0.035), rather than zero (p=0.048). Additionally, more than 80% of the data fall outside of the threshold for 'equilibrium' (±20% EPC_{sat}; Jarvie et al., 2005). In summary, we reject our first hypothesis that EPC_0 and DRP tend towards an equilibrium at zero difference. Rather, disparity between EPC_0 and DRP appeared to be the norm.

2.4.3 Does PEP vary downstream when the P inputs change?

To test our second hypothesis – that PEP is invariant for new inputs of P in the stream – we found 7 studies (n=37 pairs) with paired observations above and below a point-source. In general, these point-sources were WWTPs with concentrations of effluent DRP typically measuring several mg P L^{-1}. The median increase in DRP concentration downstream of WWTPs was 0.23 mg P L^{-1}, ranging from 0 to 6.65 mg P L^{-1}. This led to a consistent increase in EPC_0 as well (Figure 2.5), where EPC_0 was greater downstream by a median of 0.18 mg P L^{-1}; the Wilcoxon signed-rank test estimated an average increase

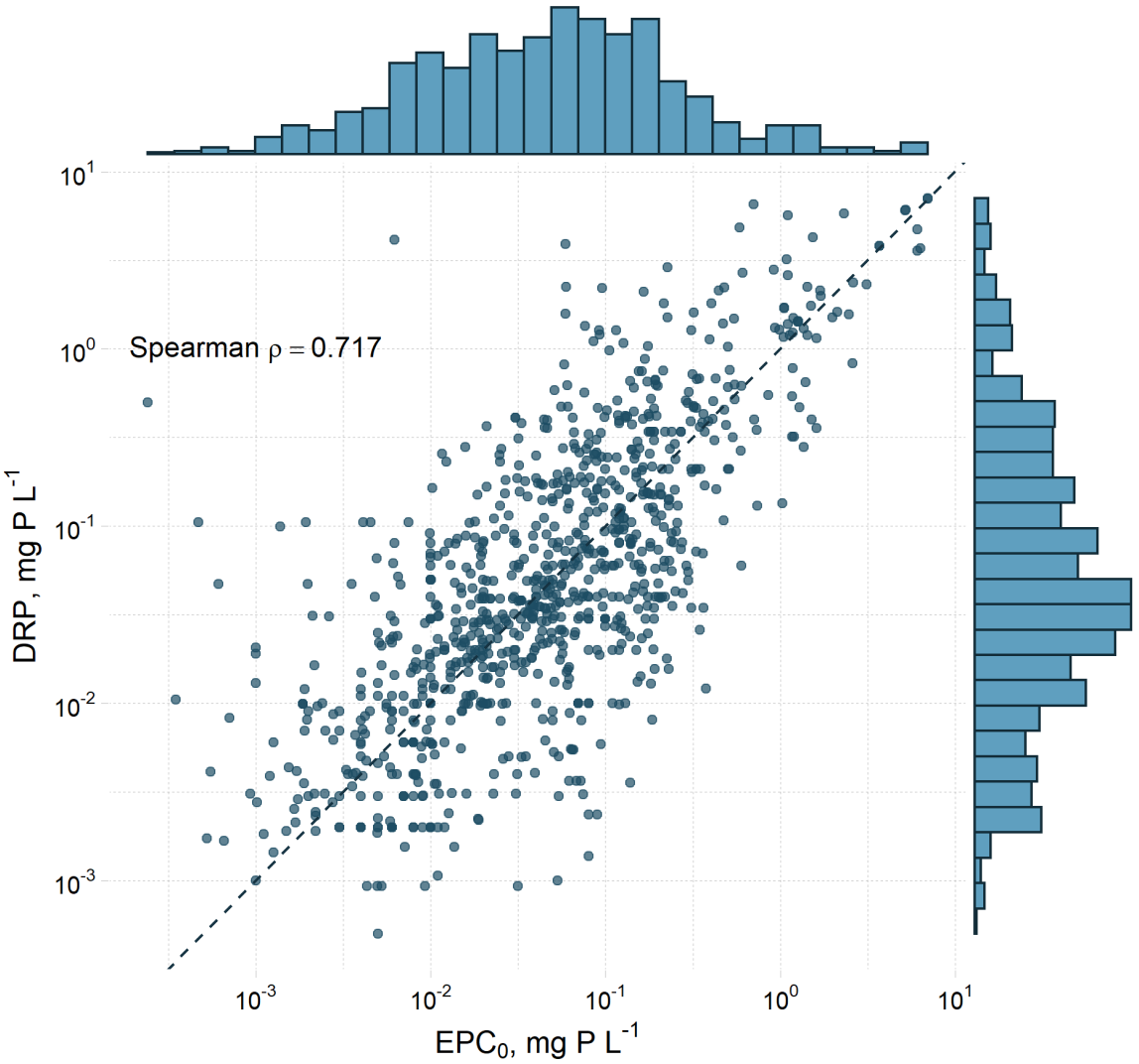

Figure 2.3 Scatter plot of paired observations of DRP and EPC_0 collected through the systematic review (*n*=942). Both variables are shown on $\log_{10}$ scale, with marginal histograms shown as well. The dashed line is the 1:1 line. DRP and EPC_0 are highly correlated (*p*=1.8e-149), as shown with the Spearman's correlation test.

of 0.95 log units across these studies (*p*=2.90e-7) . However, there was no consistent shift in PEP downstream from a point-source for these observations (*p*=0.47). These data suggest that the in-stream processes controlling any difference between DRP and EPC_0 (PEP) was not biased to a significant degree by the presence of effluent enriched in DRP. We do note that these pairs of observations covered <10% of the full dataset and the test did not account for other moderating variables such as changes in pH. We explore the influence of point sources with more variables in detail in the following section.

2.4.4 What causes the variance in PEP?

To infer which variables moderated the PEP, we used multiple imputation to account for missing data. Details on the convergence and performance of this method are given in Appendix A. While multiple imputation is a robust approach and converges well in the face of poor information (Van Buuren et al.

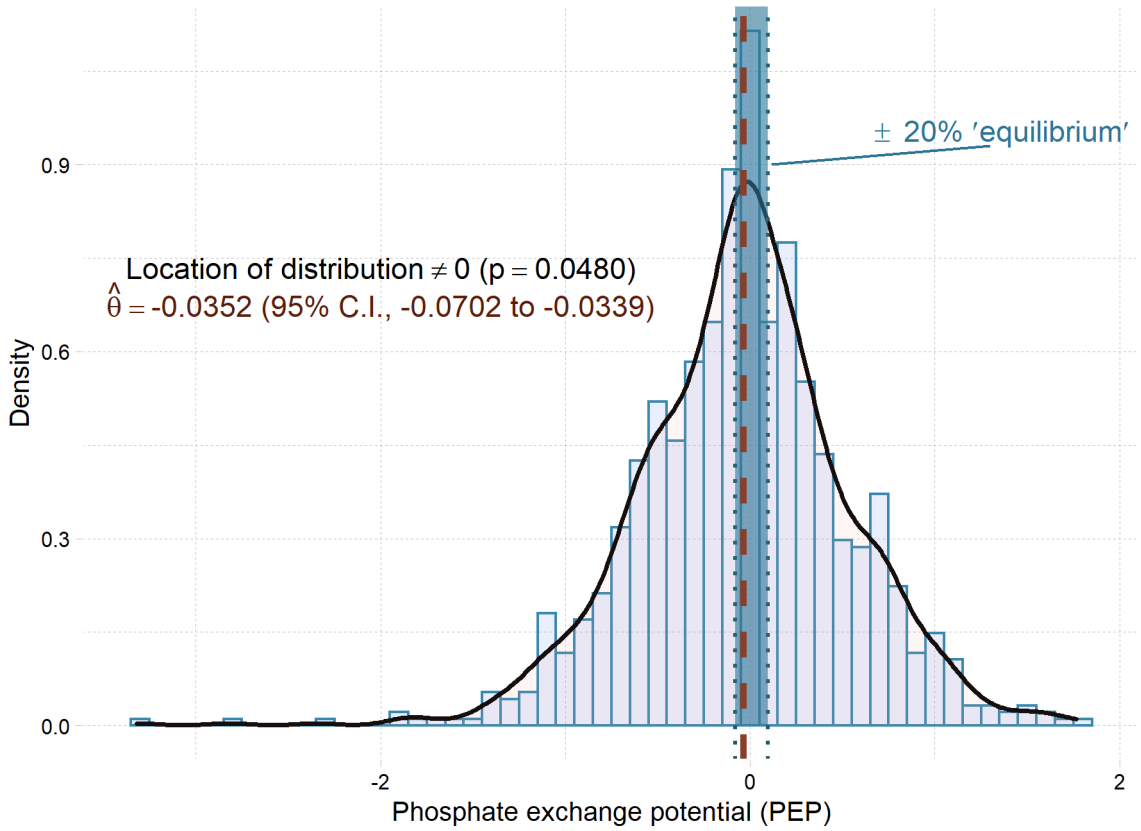

Figure 2.4 The distribution of phosphate exchange potential (PEP; $\log_{10}(EPC_0) - \log_{10}(DRP)$) across all 45 studies (n=942), where negative values (greater DRP relative to EPC_0) indicate a potential for sediments to adsorb DRP and positive values (greater EPC_0 relative to DRP) indicate a potential for sediments to release DRP. The blue shaded area indicates the ±20% EPC_{sat} threshold as defined by Jarvie et al. (2005) where, $EPC_{sat} = 100 \times [(EPC_0 - DRP)/EPC_0]$ and is transformed here to the same basis as PEP; 83% of the data fall outside of this threshold. A Wilcoxon signed-rank test for the location of the PEP distribution rejected the null that it was equal to 0; the estimated location of the PEP distribution (pseudo-median; $\hat{\theta}$) is indicated with the red dashed line.

2006), some variables (e.g., k_{Lang}) were poorly observed while others contributed more information (e.g., fines concentration): we stress that model interpretation with this much missing data still requires caution.

The multiple imputed datasets were each modelled with a GAM before being pooled to a single, final model (Figure 2.6; Table A.4). Considering the variability in PEP (standard deviation of 0.571), the individual GAM fits performed reasonably well, with average RMSE of 0.476 (±0.011, range from 0.442 to 0.502). When pooling the models, the parametric terms had between-imputation variance lower than within-imputation, so the fractions of missing information (FMI) were moderate (0.19 to 0.40), suggesting that inferences about these parameters are fairly robust. Notably, variability in EPC_0 methodology was responsible for some differences in PEP between studies. PEP decreased with longer equilibration times (-0.21 units per $\log_{10}$ h) while pre-treating sediment via freezing/freeze-drying or air-drying increased PEP by an estimated 0.36 and 0.24 units, respectively. Increasing solution ionic strength, however, had a tenuous effect on PEP, with a decrease of approximately 0.005 for a 1 mM increase in ionic strength used.

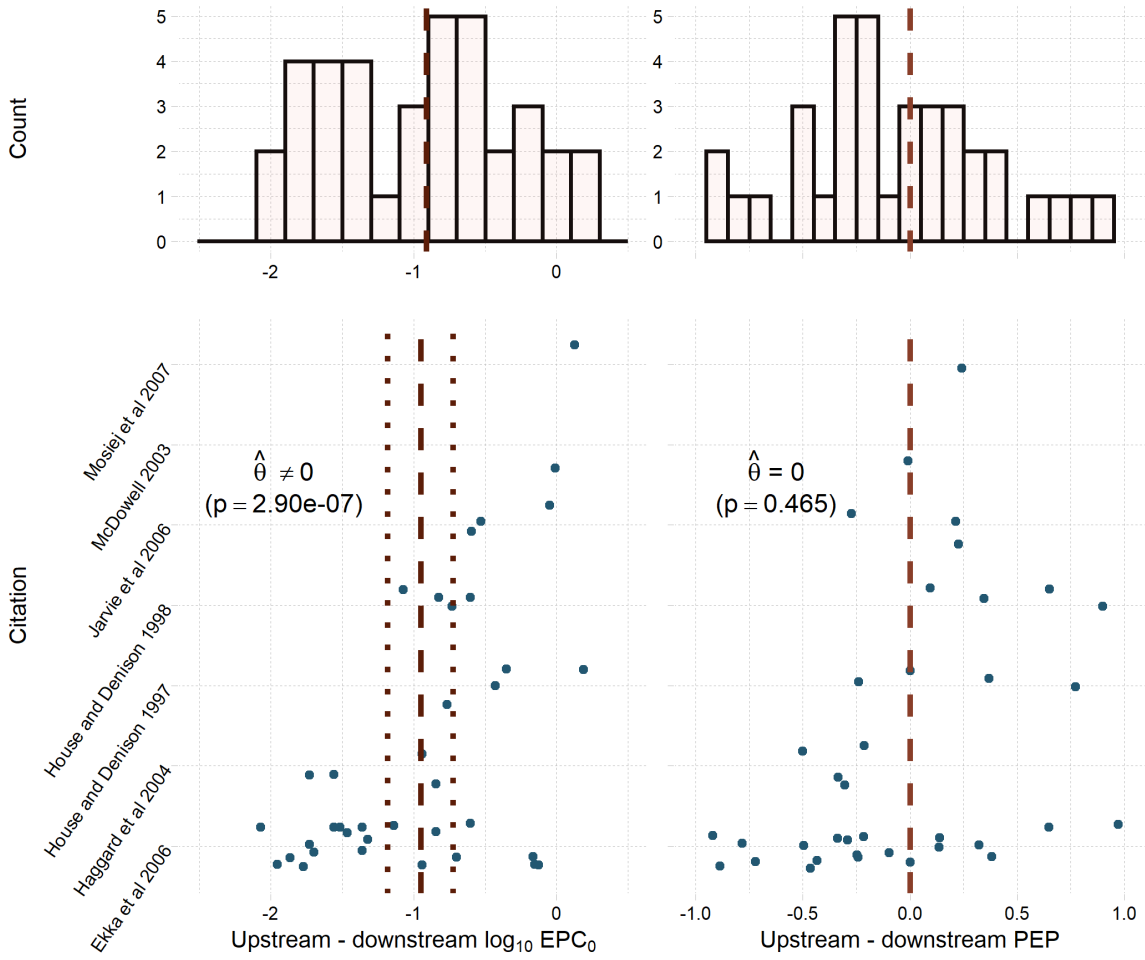

Figure 2.5 The distribution of changes in EPC_0 (left, log-transformed) and PEP (right) when measured upstream and downstream of a point source across the entire dataset given as a histogram (above) and by individual study (below). Seven studies with a total of 37 observations were used to test whether EPC_0 and PEP change downstream with a major P input (measurements were taken on the same stream on the same date above and below a reported point source). For EPC_0, a paired Wilcoxon signed rank test rejected the null hypothesis and estimated an overall change in EPC_0 of -0.95 log units when below a point source (indicated with the dashed line, dotted lines indicate the 95% confidence interval). For PEP, the test failed to reject the null hypothesis that this distribution's location was zero.

In contrast to the paired test result in section *2.4.3* (above), the pooled model here suggested that PEP decreased by 0.099 with point-source input (meaning the disparity favored a greater DRP relative to EPC_0 when downstream from point-sources). The contrasting result with section *2.4.3* is likely a function of the different data used (here, the full, multiply-imputed data and above only 7 studies) and the consideration of more variables, which captured some of the wide variability in PEP. In summary, the test on paired observations was inconclusive but our multiple imputation approach utilized more information to estimate a modest effect of point-sources (i.e., likely due to sudden increases in DRP concentration or a saturation of sorption sites) with less potential for sediment P release (lower PEP).

Smooth terms from the pooled GAM, shown in Figure 2.6, were less certain than the parametric terms since the approximated fractions of missing information ($\widetilde{FMI}$) ranged from 0.52 to 0.82 (see Table A.4). Nonetheless, these terms provide a helpful summary of how these complex variables might moderate

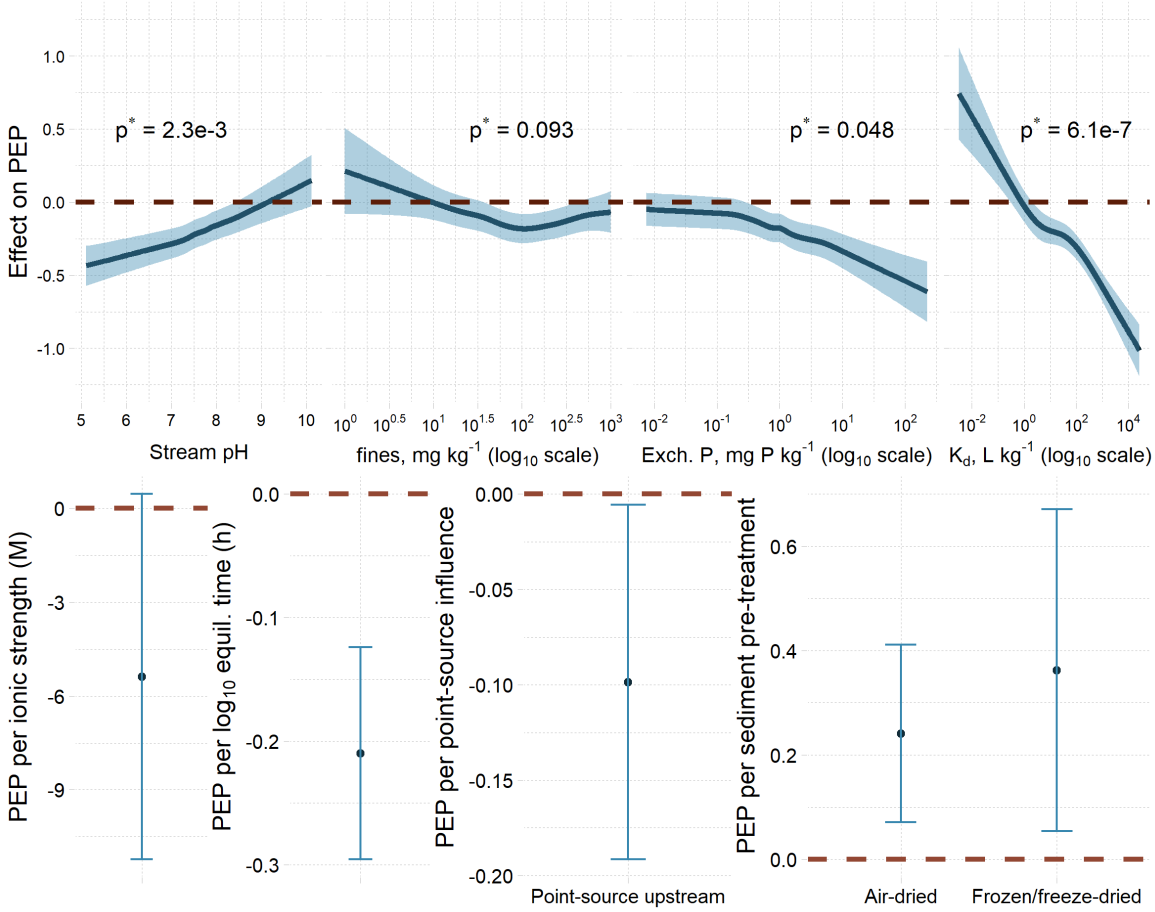

Figure 2.6 Plots of the smooth terms (top row) and parametric terms (bottom row) in the GAM fit generated by pooling the multiple imputations. The smooth additive functions are plotted by holding the other terms constant (at typical values), where the *y*-axis is the predicted *partial response* (in same units as PEP). The shaded area indicates an estimated 95% confidence interval about the fit; the approximated *p*-value for these terms (p^*) are shown to aid interpretation. The parametric (linear) terms are shown as their estimate and 95% confidence interval; note that these terms are slopes for ionic strength and $\log_{10}$ equilibration time but are intercept terms for upstream point-source and alternative sediment pre-treatments (relative to no upstream point-source and sediments analyzed fresh). Study-level random effects are not shown. For full details, refer to Table A.4.

PEP. The PEP tended to decrease as both sorption strength (via $\log_{10} K_d$) and complexed P (via $\log_{10}$ exchangeable P) increased. PEP also tended to decrease with $\log_{10}$ fines concentration, but this pattern tapered off after a few hundred g kg^{-1} of fines. The relationship with stream pH was uncertain at extreme ends where fewer data were available (pH <6 and >9), but overall suggested that PEP trended upwards with increasing pH. The random effects in the model (estimated standard deviation of 0.0074) suggested that some study-specific variance was not accounted for by the data available for the GAMs (*p*=8.7e-7), although this variance was small relative to the remaining variance (standard deviation of 0.464).

2.4.5 Does PEP vary according to site and/or by season?

Across all 45 studies, we identified those where three or more observations were recorded for separate dates at a given site, resulting in 15 studies with a total of 106 streams and 413 observations (Figure A.7). These data spanned multiple seasons and sometimes years yet showed a relatively tight variation with respect to site for both EPC_0 and for PEP compared to the whole dataset (Figure 2.4). Coefficients

of variation (CVs; dimensionless) for within-site EPC_0 had a median of 0.022 while the review-wide CV for EPC_0 was 3.1. For PEP, the within-site standard deviations had a median of 0.27 units compared to 0.57 units for the whole dataset. Little variation, however, was noted due to seasonal differences in PEP (Figure A.8). Using sites with $n \geq 5$ observations, the within-site variation generally did not correlate with time of year.

2.4.6 Do sediment microbes influence sediment P sorption?

Six studies identified in this review made comparisons between fresh sediment (microbial activity present) and sterilized/inhibited sediment (microbial activity absent) for various P sorption metrics, including EPC_0. We present the qualitative results and interpretations in Table 2.1. Early studies by Meyer (1979) and Klotz (1985) observed some apparent differences for sterilized sediments, but cell lysis may have contributed some of the P released into solution. Later studies found increases in EPC_0 following sterilization but changes in P sorption indices were inconsistent: Haggard et al. (1999) and McDaniel et al. (2009) observed decreases in P sorption indices, while Munn and Meyer (1990) and Lottig and Stanley (2007) either observed no changes in P sorption indices or an increase following $HgCl_2$ sterilization. Derived biotic contributions to P uptake varied in a similar fashion: 0 to 50%. These studies differed in their approaches to potential cell lysis and a subsequent flush of microbial biomass P; one study tested a phosphorylation inhibitor to control biotic P uptake but found P sorption indices equivalent to that of fresh sediment (Klotz, 1985). In summary, a biotic contribution to EPC_0 could be plausible in some cases, but the results here are unclear and possibly biased due to methodological choices.

2.5 Discussion

> *"In estuarine, coastal, and oceanic sciences, the phosphate buffer mechanism has come to mean the influence of sediments, whether benthic or suspended, in controlling the dissolved reactive phosphate concentration in the water at some near-constant value regardless of biological removal and input reactions. The terminology suggested an analogy to pH buffering."*
>
> *- Froelich, 1988*

2.5.1 Disparity between DRP and EPC_0 is the norm and may be site-specific

With >900 pairs of DRP and EPC_0 observations in streams at baseflow from 45 studies, we found that considerable disparity between DRP and EPC_0 in streams at baseflow is the norm and that only 17% of the data fell within a cutoff for an 'equilibrium' (*sensu* Jarvie et al., 2005). Sediments certainly have the potential to buffer P in streams ($PEP \neq 0$) and the direction and magnitude of the potential varies by stream (Figure 2.6; Figures A.7 and A.8 in Appendix A).

Table 2.1 Qualitative comparisons between fresh sediment and microbially inhibited/killed sediment P sorption indices. Techniques for a given study are the sterilization/inhibition technique applied. The effect on the P sorption metric given is compared qualitatively to the fresh control, e.g., a ↑ for EPC0 means EPC0 increased with the sterilization technique relative to the fresh sediment. Apparent conclusions are based on the hypothesis that decreased P retention in treated sediments is solely due to lack of microbial P uptake; actual conclusions in text may have differed.

Study	Technique	P sorption metric	Effect relative to control	Apparent conclusion on biotic contribution?	Notes
Meyer (1979)	Autoclave	P sorption rate	↓→	Minor	No difference after 8 h
	$HgCl_2$	P sorption rate	↓→	Minor	No difference after 8 h
	γ-irradiation	P sorption rate	→	Negligible	Equivalent to fresh throughout
	Autoclave	P desorbed in 0 P solution	↑	Significant	Attributed to cell lysis
	$HgCl_2$	P desorbed in 0 P solution	↑	Significant	Attributed to cell lysis
	γ-irradiation	P desorbed in 0 P solution	↑	Significant	Attributed to cell lysis
Klotz (1985)	Autoclave	P sorption index[#]	↓	Significant	Attributed to lysis
	CCCP[*]	P sorption index	→	Negligible	CCCP prevented lysis
Munn and Meyer (1990)	γ-irradiation	EPC_0	↑	Significant	Considered EPC_0 responsive to biotic activity; no mention of lysis
	γ-irradiation	P sorption index	→	Negligible	Differed between streams but not for microbial treatment
Haggard et al. (1999)	Autoclave	P sorption index	↓	Significant	Decreases of 28 to 50%, attributed to biotic uptake
Lottig and Stanley (2007)	$HgCl_2$	Sediment P uptake rate[§]	→↓	Significant for larger particles	Derived biotic contributions of 0 to 50%
	$HgCl_2$	P sorption index	→↑	Significant only for large gravel	Unknown why P sorption index increased for the 'killed' large gravel
	$HgCl_2$	EPC_0	↑	Significant for all particle sizes	Increased for all particle sizes but most so for the larger particles; cell lysis discussed to be not a major source of error
McDaniel et al. (2009)	Autoclave	P sorption index with lysis correction[¶]	↓	Significant	Derived biotic contributions of 26 to 40%

#P sorption index refers to any single point isotherm (typically with a large P concentration); e.g., 'PSI', the Bache-Williams Index, etc.
*CCCP is carbonylcyanide m-chlorophenylhydrazone, a phosphorylation inhibitor.
§Defined as a sorption measurement with 100 µg P L-1 solution for 24 h.
¶Correction for lysis was by determining P release on control sediments after autoclaving (no added P) and subtracting this quantity from data in the P sorption index measurements.

Repeatedly, a correlation between EPC_0 and DRP is taken as evidence for sediments moderating DRP in the water column. We also observed such a correlation (Figure 2.3) but emphasize that this is merely an effect of DRP and EPC_0 being interdependent variables: high DRP can increase the sediment EPC_0 (above and below WWTPs is a clear case, Figure 2.5) but a high EPC_0 can, in turn, increase DRP (Stutter et al., 2010). An

important distinction in the previous sentence is our use of 'can' and not 'will': there is only a *potential* for sediments to adsorb or desorb P relative to the water column. By using interdependent (rather than independent) variables such as EPC_0, our predictions of DRP in the water column will suffer from faulty inferences. Hence, we focus on the disparity between DRP and EPC_0 (i.e., PEP) and concern our discussion with why this potential is generated or, conversely, not always utilized.

The data suggested a tendency for the potential of sediments to remove DRP from the water column during baseflow (PEP of -0.035; Figure 2.4), although PEP varies widely (15% of the data show ≥0.5 order of magnitude difference). This highlights that the 'equilibrium' in EPC_0 stands not for 'equilibrium between P sorbed to sediments and that in the water column' but a quasi-equilibrium for a well-mixed batch solution. This difference is key for understanding the utility of EPC_0 in describing the stream sediment P buffer. In reality, P in the stream water column and P sorbed to benthic sediments is not at equilibrium, but – depending on kinetic factors (i.e., reaction limiting factors) and transport conditions (i.e., hydrological exchange flows; Harvey 2016) – could reach a steady state (Figure 2.7).

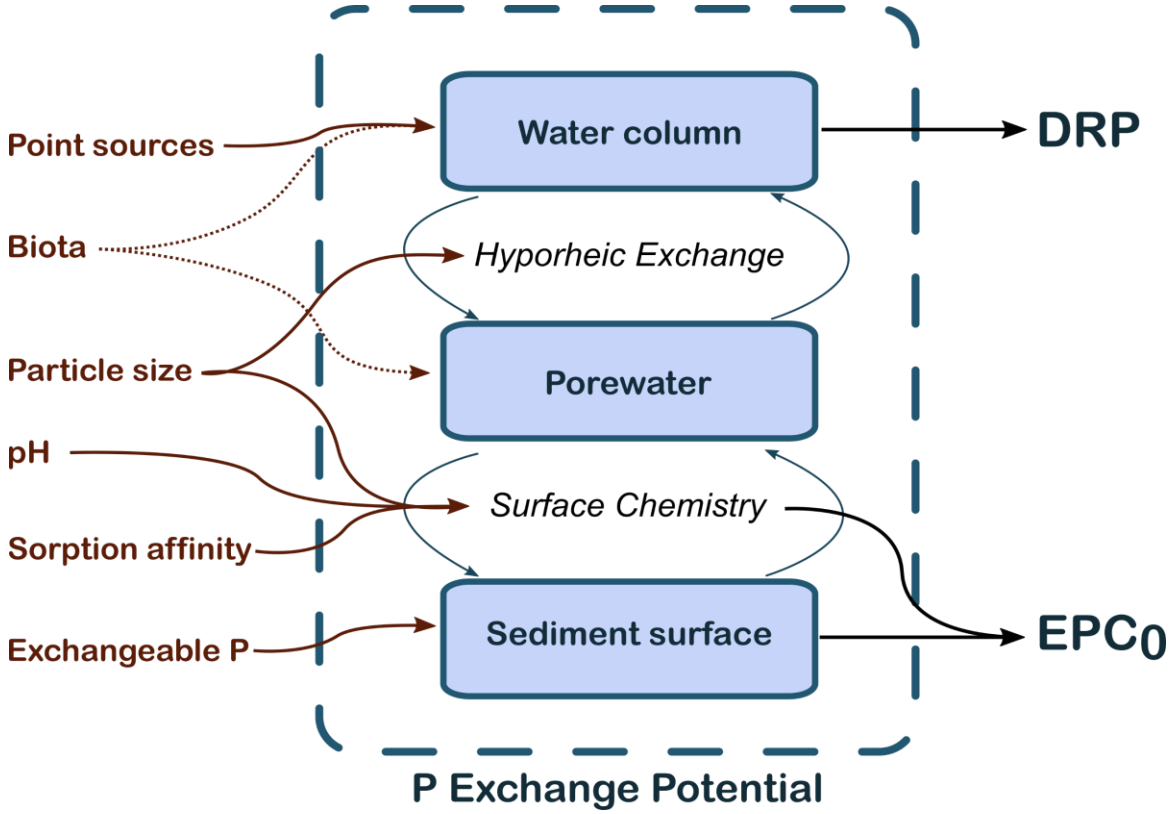

Figure 2.7 Conceptual overview of the sediment P buffer in streams. Within the dashed area, which determines the P exchange potential, boxes indicate locations where P may be at any one time: complexed with the sediment surface, dissolved within the surrounding porewaters, or in the water column of the stream. Exchange between these locations is mediated (1) by reactions at the sediment particle surface (primarily surface complexation: specific adsorption of phosphate or (co-)precipitation of phosphate) and (2) by transport between the porewaters (i.e., hyporheic zone) and the water column (via hyporheic exchange). On the left, key variables discussed in the review influence either the stock of P within the sediment P buffer (e.g., point sources can elevate DRP concentrations in the water column) or the processes that mediate exchange (e.g., sediment particle sizes influence hyporheic exchange). Note that biota may not strictly be part of the sediment P buffer but can influence the P available in either the water column or porewater. Further, the EPC_0 is a function of both the mass of P complexed with the sediment surface and the surface chemistry underlying the sediment P sorption strength.

Below, we synthesize our meta-analysis on EPC_0 to discuss what influences the disparity between EPC_0 and DRP (PEP) by acting on both variables and EPC_0 alone.

2.5.2 Influences on EPC_0 and DRP simultaneously

2.5.2.1 Water-column pH

As stream pH increased, so did PEP (Figure 2.6). Being a master variable for in-stream biogeochemistry (Wetzel 2001; Nimick et al. 2011), mineral equilibria (Lindsay et al. 1989; Sø et al. 2011), and sorption chemistry (Goldberg and Sposito 1985; Arai and Sparks 2007), pH has a strong yet complex influence on PEP, as it can moderate both DRP and EPC_0. We highlight just the most prominent influences on PEP.

In this review, reported stream pH ranged from 5.2 to 10.1 while sediment pH ranged from 4.8 to 8.6. So, $H_2PO_4^-$ likely dominated the total phosphate speciation in these waters (Lindsay et al. 1989; Stumm and Morgan 1996). At lower pH, however, metal oxide surfaces on sediments exhibit greater phosphate sorption strength (Sigg and Stumm 1981; Goldberg and Sposito 1984a; Strauss et al. 1997; Zhou et al. 2005; Arai and Sparks 2007). Hence, low pH likely favors strong P sorption rates thus driving EPC_0 down (Meyer 1979; Klotz 1988). For example, above and below acid-mine drainage inputs in four streams, Simmons (2010) observed a stark increase in sediment sorption affinity (as K_d) and a decrease in EPC_0 concordant with the ~2-3 unit drop in pH. Depending on the stream hydrology, however, there may be insufficient time (or space) for water-column DRP to contact the sediment surfaces, thus causing a negative PEP.

As pH increases, processes separate from the sediment surface may remove DRP from the water column. For example, (co-)precipitation with Ca minerals may remove phosphate from solution in calcareous streams (House 2003; Cohen et al. 2013; Corman et al. 2016). Further, a higher pH may indicate greater autotrophic activity (due to CO_2 removal), where periphyton may assimilate DRP according to stoichiometric demand (Mulholland et al. 1994; Dodds 2003; Hill et al. 2012). While these processes can lower DRP coincident with greater pH, they may not be reflected in sediment EPC_0, hence the tendency for greater PEP at higher pH.

2.5.2.2 Benthic sediment fines concentration

Like pH, benthic sediment fines concentration simultaneously affects EPC_0 and its relationship with DRP. We found that PEP initially decreased with an increasing concentration of fines, but after approximately 100 mg kg^{-1}, the relationship flattened out (Figure 2.6). The concentration of fines likely alters PEP by controlling two principal variables: the number of sorption sites and the degree of contact between sorption sites and the water column.

Finer sediments often have greater P sorption affinity than coarser sediments (Stone et al. 1995; McDowell et al. 2003; Haggard and Sharpley 2007; Lottig and Stanley 2007; McDaniel et al. 2009; Agudelo et al. 2011) as finer sediments generally have greater specific surface areas, more clay minerals

(Gérard 2016), and more metal oxides (Rawlins 2011) and thus more P sorption sites. As a result, greater P loading is required to raise the EPC_0 of finer sediments in comparison to coarse sediments (see section *2.5.3.1.*). Hence, lower EPC_0 occurs more frequently for finer sediment (Haggard et al., 2007; McDaniel et al., 2009; Agudelo et al., 2011; McDowell, 2015), which may explain the initial declining trend in PEP from 0 to ~100 mg kg^{-1} fines. A counter process to this pattern may be the mineralization of organic P compounds contained in fine, organic-rich sediments (Tank et al. 2010; Baldwin 2013; McDowell and Hill 2015), which, when degraded, may release mineralized P to porewaters (Klotz 1985; McDaniel et al. 2009; Weigelhofer et al. 2018a). The contribution of sediment organic P mineralization to P exchange in streams is poorly understood, but it likely depends on the hyporheic microbial community (Findlay 2016; Battin et al. 2016), nutrient availability (Sinsabaugh et al. 2012), and substrate quality (Cross et al. 2005).

While finer sediments are generally more sorptive, P exchange in streams will also depend on hydrologic (i.e., hyporheic) exchange (Figure 2.7). As a batch incubation method, EPC_0 employs ideal contact between solution and sediment particles during shaking whereas, in streams, restricted contact occurs between the water column and sediment surfaces as mediated by hyporheic exchange – that is, vertical hyporheic exchange flows are often orders of magnitude lower than the stream flow (Boano et al. 2014; Hartwig and Borchardt 2015; Harvey 2016). Despite the potential for P sorption, fine sediments restrict advective flows through the streambed while dispersive transport through the benthic substrate is often insignificant relative to the stream's hydraulic load (Aubeneau et al. 2014; Weigelhofer et al. 2018a). Indeed, stream P uptake can increase with greater hyporheic exchange (Orr et al. 2009) and volume of transient storage (Ensign and Doyle 2006; Bohrman and Strauss 2018), thus leading to reduced P uptake in streams whose geomorphology is limited in these hydrological characteristics (Weigelhofer 2017; Booman and Laterra 2019).

To illustrate, consider the two sites on the forested, Hoxie Gorge stream in Klotz (1991) and shown in Figure A.7: there is a clear difference in both EPC_0 and PEP between the sites. The upper site (upstream of a beaver dam, providing slow flow and fine, reactive sediment) consistently maintained a lower PEP than the downstream site on the same stream (below the beaver dam, coarse substrate with fast, turbulent flow), meaning the finer reactive sediment upstream had greater potential to remove DRP and indeed may have decreased DRP for the downstream site. Overall, as sediments become finer, there is a tradeoff for the sediment P buffer in lotic systems: finer sediments entail greater P sorption affinity yet reduced hyporheic exchange. This may partly explain the tendency for negative PEP (potential for DRP removal) observed across all data reviewed here (Figure 2.4). Changing geomorphology and fines concentration in benthic sediments can influence the sediment P buffer in complex ways and will require further research that connects sediment P affinity and hydrological exchange.

2.5.3 Influences on EPC_0 alone

2.5.3.1 The EPC_0-K_d-Exchangeable P triad

Aside from methodological variables (see below), two variables in our meta-analysis moderate the PEP by interacting with only the EPC_0: sediment exchangeable P concentration and K_d (a proxy for P sorption affinity). Being intimately tied not only to EPC_0 but to each other, we discuss both exchangeable P and K_d together.

The sediment EPC_0 is the concentration along a P adsorption isotherm (or buffer diagram) where *net* P sorption is zero. To increase the equilibrium P concentration of the batch solution, greater additions of P will be required for sediments with greater P sorption capacity as measured via K_d (*sensu* Froelich, 1988) or related sorption metrics, e.g., the Bache-Williams index (Bache and Williams 1971) and anion storage capacity (Saunders 1965). Likewise, for an equivalent EPC_0, a sediment with greater K_d has greater amounts of P sorbed with its surface – exchangeable P serves as a proxy for this sorbed P, and is related to other variables like native adsorbed P (Nair et al. 1984), bioavailable P via Fe strip (Sharpley 1993), and labile P fractions (Condron and Newman 2011). We note the term 'exchangeable' is difficult, though, since (1) exchangeable P is defined by the extractant used and (2) P desorption hysteresis prevents positive and negative P exchange from being symmetric (Barrow 1983b; Lair et al. 2009; Krumina et al. 2016). Additionally, sediment P sorption capacity is finite since sorption sites can become saturated (Sposito 2004). So, with greater P adsorption, the local linear approximation of the sorption slope (K_d) can only decrease; meanwhile, the exchangeable P pool will increase assuming no major 'fixation' of sorbed P occurs (i.e., made non-exchangeable or occluded; Strauss et al. 1997; Gustafsson et al. 2012). While sediments can have similar EPC_0 but very different K_d, the two variables tend to co-vary, owing to broad similarities across catchments in sorption sites on sediment surfaces (primarily Fe and Al oxides). Hence, our meta-analysis observed a sharp decrease in PEP with increasing K_d (Figure 2.6): highly sorptive sediments drove EPC_0 down relative to the water-column DRP. The additive component in our model for exchangeable P, however, illustrates a gradual saturation of sorption sites with increasing exchangeable P (meaning less ability to buffer DRP), which can be observed in streams with very high DRP. The saturation of sorption sites may also explain the significant decrease in PEP downstream from point-sources, as measured by our model (decrease of 0.099).

This inter-relation between sediment EPC_0, exchangeable P, and K_d is important to sediment P buffering in many environments. Using numerous sediments across Florida Bay, Zhang and Huang (2007) showed an exponential increase in EPC_0 as a function of exchangeable P; simultaneously, K_d decreased and, further, was a function of surface reactive Fe oxides unless exchangeable P was saturating these sorption sites (exchangeable P greater than ~3.7 mg P kg^{-1}). Analogously, Haggard (2007) and Stutter and Lumsdon (2008) measured exponential decreases in EPC_0 with increasing K_d. Across a gradient of non-point source P inputs in an agricultural ditch network, Ezzati et al. (2020) measured a gradual increase in ditch sediment EPC_0 and exchangeable P (as Mehlich-3 P) while K_d decreased. Likewise, sediments below a WWTP increased in both EPC_0 and exchangeable P with longer exposure to the high DRP

concentrations in the water column (House and Denison 1997; Haggard et al. 2004; Jarvie et al. 2005; Ekka et al. 2006; Jarvie et al. 2006b; Stutter et al. 2010) which can reduce P uptake at the stream scale (Marti et al. 2004).

As shown by these observations, careful consideration of the sediment EPC_0-K_d-exchangeable P inter-relationship described here is necessary to understand the stream sediment P buffer and predict how it regulates P fluxes. To illustrate, consider the hypothetical P fluxes in three different streams at steady-state conditions shown in Figure 2.8. A high K_d entails a strong sediment P buffer which can greatly attenuate either temporary (Figure 2.8a) or sustained (Figure 2.8b) inputs of P. Note that Figure 2.8a resembles the gradual movement of legacy P observed in large catchments (Powers et al. 2016), suggesting a role for the sediment P buffer at the catchment scale. However, because a strong sediment buffer can transiently store more exchangeable P, it can sustain greater P release for longer periods of time after P inputs are mitigated. The saturation of P sorption sites results in a *diminishing* sediment P buffer (see 'weak P buffer' in Figure 2.8) but can be reflected by different EPC_0 for different sediments: information contained in exchangeable P and sorption affinity is also required for understanding how sediments may buffer P. Acknowledging this dynamic can change our interpretations, for example, of strong negative PEP (a *potential* for adsorption at equilibrium) from 'sediments are acting as sinks' to 'the sediment P buffer is likely overwhelmed', and improve in-stream P cycling models that currently rely on EPC_0 (White et al. 2014; Jackson-Blake et al. 2015).

2.5.3.2 Methodological influences on EPC_0

No standard method exists to measure EPC_0. In fact, a standard EPC_0 method may be unrealistic (Nair et al. 1984). The challenge lies in our goal of making measurements comparable between laboratories yet relevant to the stream conditions. Our analysis showed that methodological choices influenced EPC_0 measurements and so PEP as well. We discuss each of these effects here and refer the interested reader to Appendix A for a recommended baseline EPC_0 methodology.

The solution used in the batch experiments varied, with studies using filtered stream or ground waters, various synthetic solutions mimicking ambient stream chemistry, and sometimes only deionized water. We could not account for every aspect of these solutions regarding P sorption: for example, pH (Meyer 1979; Dzombak and Morel 1987; Klotz 1988) and solution Ca concentration (Nair et al. 1984; Klotz 1991; House and Denison 2000; Lucci et al. 2010) are known to affect P sorption. Nonetheless, there was possibly a decrease in PEP with an increase in solution ionic strength (Figure 2.6). This is consistent with the observation that greater ionic strength promotes P sorption and inhibits desorption (Ryden and Syers 1975; Barrow and Shaw 1979; Froelich 1988), therefore lowering EPC_0.

Equilibration time for the EPC_0 measurement also influenced PEP, as longer times favored lower EPC_0 and hence lower PEP. The majority of P sorption is relatively fast, often with exponential decay of sorption through time (McDowell and Sharpley 2003b; Arai and Sparks 2007); however, the reaction continues for unknown periods on the order of days or longer (Bolan et al. 1985; Goldberg and Sposito

1985; van der Zee et al. 1989), referred to as two-step kinetics by Froelich (1988). Any P concentration measured during a sorption experiment is therefore a concentration at a quasi-equilibrium (Barrow 1983a). While the nature of the slower reaction is unclear (but see Khare et al. 2005; Arai and Sparks 2007; Krumina et al. 2016), the faster kinetics are the most critical for the sediment P buffer. An appropriate equilibration time for EPC_0 needs to be a tradeoff between stability of measurements (i.e., by reaching a sufficient quasi-equilibrium) and relevance to the timescales for transport within the stream (e.g., hyporheic zone residence times can span seconds to days; Boano et al., 2014). An overnight (e.g. 16 or 24 h) shaking may be sufficiently stable, practical, and reasonably reflective of the equilibrium condition of the sediment surface.

The last methodological variable for EPC_0 we examined was the sediment pre-treatment (i.e., drying). Relative to measurements on fresh (wet) sediments, we estimated increases in PEP of 0.36 and 0.24 for analyzing sediments frozen/freeze-dried and air-dried, respectively. It should be noted, though, that 92% of the observations used fresh sediments. Nonetheless, this finding supports the recommendation of fresh sediments over drying pre-treatments (Klotz and Linn 2001; Attygalla et al. 2016; Simpson et al. 2019). Drying (especially air-drying) alters sediment biogeochemistry in complex ways. Particularly, it: lyses sediment microbial P (Qiu and McComb 1995; Turner and Haygarth 2001), alters sediment organic matter (Turner et al. 2007), oxidizes Fe^{2+} and ages hydrous metal oxides responsible for P sorption (Phillips and Lovley 1987; Baldwin 1996; Qiu and McComb 2002; Attygalla et al. 2016). Possible storage artifacts for fresh sediment EPC_0 due to microbial activity are not well known but could conceivably increase EPC_0 by lowering the redox potential (thus mobilizing some Fe-bound P). However, Rahutomo et al. (2018) demonstrated either no or little increase in EPC_0 when sediments were stored under anaerobic conditions for 30 days; further, the redox effect was magnified only with the addition of labile carbon, illustrating the fact that carbon limitation – which is prevalent among stream sediments (Hill et al. 2012) – likely limits the decrease in redox potential necessary to bias EPC_0 on reasonable timescales. Hence, the bias due to pre-treatment can and should be avoided by analyzing sediments fresh as soon as possible and within 1 to 2 weeks of sampling.

Several other methodological variables must be considered when measuring and interpreting EPC_0. Redox status should be considered when planning benthic sediment sampling because it influences EPC_0 and its relation to DRP (Palmer-Felgate et al. 2011; McDowell et al. 2020a). If the goal is to characterize P exchange between the water-column and the benthic sediments at the sediment-water interface (Figure 2.7), then sampling should avoid sediments deeper in the streambed or in stagnant zones that not only interact minimally with the water-column but are also more likely to be anoxic. Incubation temperatures must also be comparable (whether within a given stream or between streams), as greater temperatures can increase the EPC_0 – a result known from surface chemistry work (Barrow 1983a; Goldberg and Sposito 1984b; Arai and Sparks 2007) and observed in empirical research (Barrow and Shaw 1979; Klotz 1991; McDowell et al. 2017). The pH of the solution has a dramatic effect on P sorption by both altering the ligand exchange mechanism for hydrous metal oxides (Sigg and Stumm 1981; Dzombak

and Morel 1987) and by altering mineral equilibria (Lindsay et al. 1989; Stumm and Morgan 1996). Future work could examine whether a weak Tris buffer or similar adjustment at circumneutral pH could standardize the pH and prevent effects of pH change during the incubation (Meyer, 1979; Klotz, 1991).

Appropriate P concentrations and careful statistical estimation of EPC_0 (the *x*-intercept in a buffer diagram) are pivotal to accurate measurement of EPC_0 (Simpson et al., 2019). Regression is fallible, and the assumption of linear sorption is increasingly violated at concentrations further from the EPC_0, leading to biased estimates (Froelich, 1988). For example, when the lowest non-zero P concentrations used for the batch sorption experiments are too high, the resulting regression estimates can sometimes yield negative values (Smith et al. 2006; Smith 2009). As negative EPC_0 is physically impossible, this highlights that care is needed when selecting initial P concentrations, particularly when the sediment has weak P sorption capacity, and when fitting isotherms. Ambient concentrations with points near the true EPC_0 (e.g., ~10 to 100 µg P L^{-1} for most sediments reviewed here; Table A.2) will give the best results while points further away will only diminish the accuracy. This could be a limitation for the use of environmental waters as the background solution, since ambient DRP concentrations may preclude desorption points (cf. Henson et al. 2019). On the other hand, in cases of severe P pollution (e.g., downstream of large point-source), greater P concentrations (>1 mg P L^{-1}) will be necessary to measure adsorption points for EPC_0 (Roberts and Cooper 2018).

As a final point on EPC_0 methodology, we generally do not recommend using environmental waters (e.g., filtered stream water) as the background solution for three reasons. First, solution composition will vary between and even within studies and hence is not replicable. Relatedly, due to logistical constraints, streamwater composition (e.g., pH) can change significantly between time of sampling and time of EPC_0 measurement. Second, ambient DRP concentrations may be too high and could preclude measurable desorption points. Third, ambient DRP concentrations in environmental waters are not equivalent to phosphate concentrations: we assume that the reacting species in P sorption assays is primarily orthophosphate but DRP can include other P species (e.g., labile organic or colloid-bound species) that are detected by the common molybdenum-blue method (Haygarth et al. 1997; Nagul et al. 2015; Maruo et al. 2016; Worsfold et al. 2016). Hence, in waters where these interferences are likely (e.g., Jansson et al. 2012), the part of DRP that is not phosphate can positively bias the EPC_0.

In summary, no perfect method for EPC_0 exists (Nair et al., 1984). However we can make our measurements robust and representative to maximize inter-laboratory consistency and relevance to the stream. To this end, we have detailed an approach in Appendix A. Further, we note that EPC_0 only attempts to measure the 'level' of the sediment P buffer (Froelich, 1988), which is determined by the history of P loading to that sediment and its P sorption capacity (see section *2.5.3.1.*). So, we also recommend measuring relevant sediment P pools (see Condron and Newman, 2011) and P sorption capacity (e.g., K_d) alongside EPC_0. Together, these three sediment P metrics will characterize sediment P status more comprehensively than any one metric alone (Froelich, 1988).

2.5.4 Towards an improved understanding of the stream sediment P buffer and EPC_0's utility

The sediment P buffer is critical to understanding P transport in streams. Sediments can attenuate P loads, thus dampening the intensity but, through subsequent P release, can prolong the duration of P fluxes to downstream ecosystems: the stream sediment P buffer is part of the legacy P phenomenon (Sharpley et al. 2013). Again, consider the three hypothetical streams shown in Figure 2.8a (assuming steady-state conditions). In one, the sediment P buffer is weak and so the P flux reaches downstream relatively quickly and at greater instantaneous rates. However, if the sediment P buffer is stronger, this flux is attenuated so that rates are dampened but the delivery is elongated; in other words, the stronger sediment P buffer 'flattens the curve' of P flux downstream. Note, the total P flux downstream integrated over time (total mass) is equal in both these cases. In a third case, the stream has a strong P buffer as well as long-term P removal mechanisms, which may be related to geochemical (e.g., P fixation to recalcitrant forms) or biotic (e.g., stored in recalcitrant organic matter) processes; here, the total flux downstream is diminished and the sediment P buffer serves to give greater opportunity for this P sequestration.

Attenuation of P inputs is, on one hand, beneficial as it provides greater opportunity for P removal (in long-term storages; Figure 2.8). On the other hand, it obscures current P sources and stocks within the catchment as part of legacy P, making changes to P management challenging to evaluate (Meals et al. 2010; Hamilton 2012; Powers et al. 2016; Macintosh et al. 2018). Partly due to this lag in P transport, best management practices can take years or more to yield reductions in P losses from non-point sources (Meals et al., 2010). In fact, like in Figure 2.8b, streams with a strong sediment P buffer and/or little disturbance of the streambed may take much longer to respond to reductions in sustained P inputs (Jarvie et al. 2013b). The sediment P buffer is key to understanding this P attenuation and, therefore, improving our management for water quality.

We know that EPC_0 and DRP are positively correlated, and that 'disequilibrium' is the norm. This implies that, in most situations, knowing EPC_0 will give an indication of the potential direction of the sediment P buffer. These data are useful if, for example, one wants to know how the sediment may potentially influence water-column DRP concentrations once inputs from WWTPs cease. Or, for another example, how newly deposited sediments may affect DRP following changes in P management on land

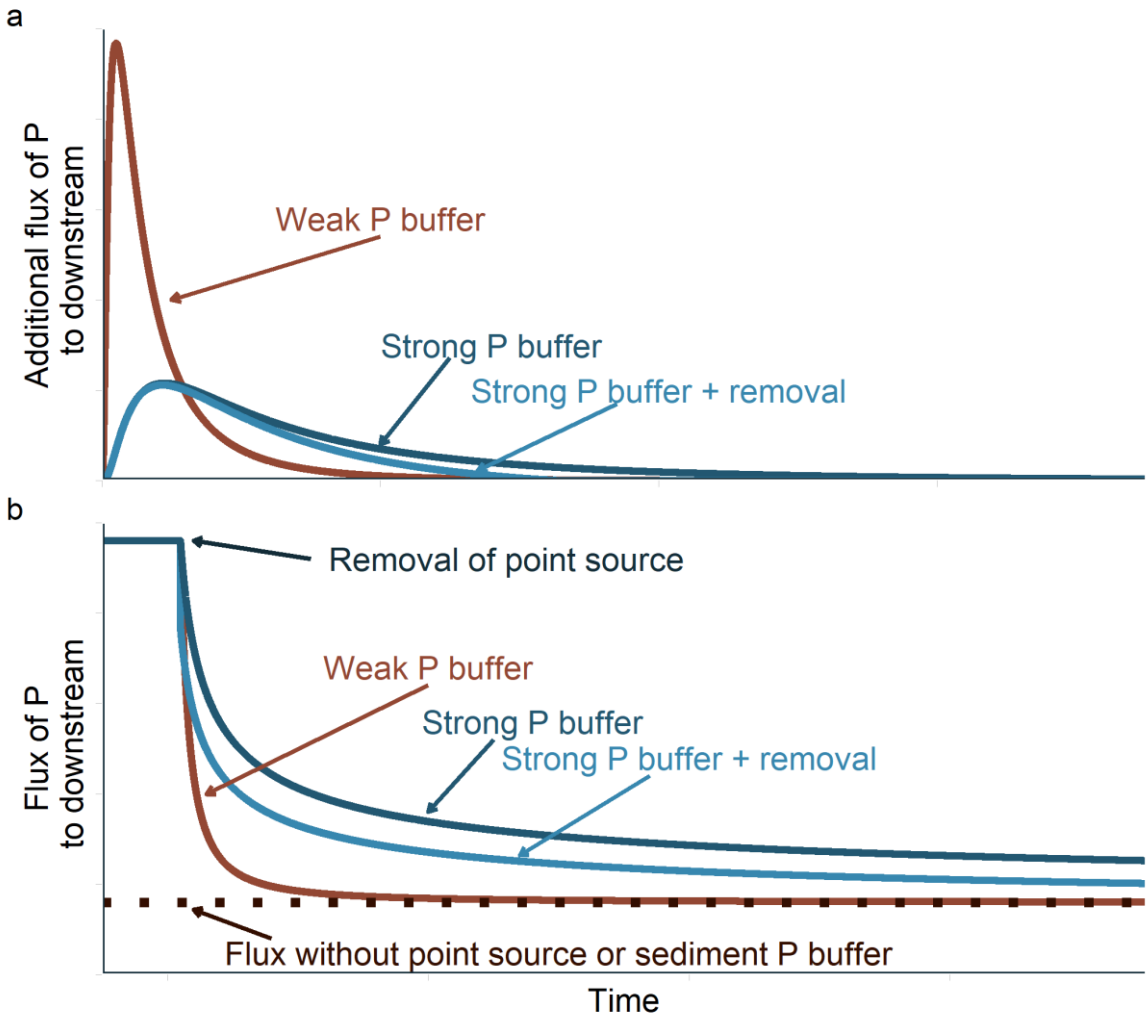

Figure 2.8 Conceptual relationships between stream P fluxes and the sediment P buffer for three hypothetical streams at steady-state conditions but for two different P input events. A 'pulse' input of P (a) to the system is attenuated -- the streams differ in the strength of the sediment P buffer and whether there is long-term removal of P within the stream. Note that the integrated P flux (or cumulative load) is equal between the weakly and strongly buffered streams. In another scenario (b), a long-term point-source of P dramatically reduces its P load to the stream – after the long exposure, all sediments would have similar EPC_0. However, the stream with the strong sediment P buffer maintains greater P fluxes in the stream and for a longer time whereas the weak sediment P buffer quickly returns to a new baseline. Log-normal distribution functions (a) and exponential functions (b) are used here to approximate the adsorption-desorption hysteresis pattern in sediment P sorption and the effects of transient storage of water within the stream. In both scenarios, significant long-term P removal (e.g., P fixation in 'occluded' form) may further reduce the P flux, particularly when a strong sediment P buffer provides opportunity for such reactions to occur.

(Macintosh et al. 2018). Indeed, improvements downstream may take a long time to manifest (Palmer-Felgate et al. 2010; Scott et al. 2011; Fanelli et al. 2019), since changes in sediment EPC_0 could take months or longer (Jarvie et al. 2006a, 2013a). Storm events may speed up this change (Jarvie et al. 2005; McDaniel et al. 2009; McDowell et al. 2019b), though further research is needed to resolve the large temporal variability in PEP (Figure A.8). We do not dispute the utility of the EPC_0 in this situation. However, EPC_0 says nothing about the rates of P release or uptake, as this is also a function of hyporheic exchange and water residence timescales in streams (e.g., as measured by the Damköhler number; Ocampo et al. 2006). While water residence times in the water column and hyporheic zone are increasingly understood and predicted (Boano et al. 2014; Ward and Packman 2019), the gap in

knowledge on sediment P exchange kinetics underlies our lack of understanding of catchment-scale P attenuation.

Recent research has suggested that the sediment P buffer comprises abiotic processes (Oviedo-Vargas et al. 2013; Griffiths and Johnson 2018; Martí et al. 2020), but biotic mechanisms for P attenuation (i.e., periphyton and sediment microbial P uptake) can be important in many streams (Mulholland et al. 1997; Dodds 2003; Ensign and Doyle 2006; Hill et al. 2012). From our review, it is unclear whether sediment biota influence the EPC_0 (Table 2.1): methodological choices seem to bias the outcome significantly due to physicochemical changes in the sediment and/or the flush of microbial biomass P upon cell lysis (Klotz 1985). Simpson et al. (2020) demonstrated for 12 streams draining a basalt geology but a variety of land uses that sediment microbes, even with likely carbon and nitrogen limitation eliminated, minimally influenced EPC_0 through biotic P demand. Importantly, though, the sediment P buffer likely moderates P supply for microbes within the benthic substrate and therefore the sediment microbial biomass P pool (Figure 2.7; Sinsabaugh et al. 2012; Weigelhofer 2017; Simpson et al. 2020). Understanding both abiotic and biotic aspects of P attenuation is important since they respond to different drivers; for example, sunlight promotes periphyton P uptake (Dodds 2003; Rier et al. 2014) but matters little to sediment P sorption. Disentangling these drivers of P attenuation is an ongoing challenge (Stutter et al. 2010; Weigelhofer et al. 2018a; Griffiths and Johnson 2018) and remains a criticism of nutrient spiraling studies (Demars 2008; Harvey 2016). However, there is hope, as when each individual process is better understood, the remaining P attenuation signal can be broken down to better understand impacts on water quality.

The sediment P buffer is dependent on the physicochemical characteristics of the sediment (particularly particle size) and so is intimately tied to the stream's sediment transport regime. As discussed earlier, EPC_0 may decrease with finer, more reactive sediments and so decrease PEP; at the same time, finer sediments constrict interaction with the water column (Figure 2.7). Benthic sediments, though, are always changing (Dietrich 1989; Hamilton 2012; Wohl 2015), meaning the streams in our idealized example (Figure 2.8) are much more variable in reality. For example, changing land uses, catchment-scale disturbances (e.g., wildfire), and stream impoundments alter sediment supply (Owens et al. 2005; Son et al. 2015; Emelko et al. 2016), which is further mediated by the geomorphological characteristics of the stream (Church 2002; Julian et al. 2016). Storm events can scour benthic substrate and deposit new material, which can change the EPC_0 (Son et al. 2015; McDowell et al. 2019b). With predictions of changing precipitation patterns (Westra et al. 2014) and more frequent floods (Hirabayashi et al. 2013; Mallakpour and Villarini 2015) for some regions under a changing climate, tying the sediment transport regime to the sediment P buffer will be critical to understand the effects on stream P attenuation.

Another important process to consider for the sediment P buffer is redox. Much of the P stored in sediments is with varying Fe species (adsorbed and precipitated) which are sensitive to reductive

dissolution (Hoffman et al. 2009; Lewandowski and Nützmann 2010; Machesky et al. 2010; Peryer-Fursdon et al. 2014; Weigelhofer 2017; Casillas-Ituarte et al. 2020). Indeed, in slow lotic systems (e.g., riverine wetlands) or in parts of the stream channel with long residence times (e.g., deeper subsurface), there tends to be greater reducing conditions which can influence EPC_0 (House and Denison 2000; Palmer-Felgate et al. 2011; Rahutomo et al. 2018). Consequently, this may lead to greater decoupling of DRP and EPC_0 *in-situ* (McDowell et al. 2020a). While the redox interface in sediments is usually below where the majority of hyporheic exchange takes place for many streams (i.e., with dissolved oxygen approximately near saturation in the water column and moderate exchange with porewaters; Boano et al., 2014), reducing conditions in zones of longer retention times or where ecosystem respiration is high (Briggs et al. 2015) may nullify the sediment P buffer (Lewandowski and Nützmann 2010; Smolders et al. 2017; Parsons et al. 2017).

2.6 Conclusions

Sediment EPC_0 and streamwater DRP indicate the potential of P to exchange between sediments and the water column. The difference in magnitudes of EPC_0 and DRP, termed the phosphate exchange potential (PEP), can be used to help set priorities and timelines for management to decrease P losses to streams. Our systematic review of the literature showed that disparity between sediments (EPC_0) and the water column (DRP) is the norm: >80% of the data showed either a significant positive or negative potential for P exchange. On average, there was a net negative PEP, or a potential for sediments to remove P from the water column. Factors that affected the direction and magnitude of the PEP were (1) chemical, for which we identified pH, K_d (sorption affinity), and exchangeable P as main factors, and (2) physical, such as particle size, which mediates P sorption affinity and hyporheic exchange, both of which can interact to influence PEP. Other factors like biotic P immobilization/mineralization and redox status may also influence the PEP, but too few data were available to clearly identify their influence.

We argue that this potential – PEP – does not translate into sediments acting as either a source or sink for P. Rather, the sediment P buffer is transient and, in addition, matters little unless there is exchange of waters between the water column and the sediment porewaters. Therefore, we emphasize that the key to understanding any kind of potential is knowing when that potential can be *realized*. With all else held equal and with no other P inputs, sufficient time and space within the stream are needed for PEP to approach zero.

More work is needed to better understand how hydrological and biogeochemical factors influence stream P cycling. This should include analyzing how variability in EPC_0, sorption affinity, and exchangeable P within and among streams is affected by environmental factors. Likewise, improved methodology for EPC_0 is needed to make inter-site comparisons more accurate while maintaining relevance to *in situ* conditions. Connecting the sediment P buffer to hydrological and biogeochemical processes in the stream will improve our predictions of how streams regulate baseflow P concentrations and loads and thereby help us manage eutrophication downstream.

Chapter 3

The error in stream sediment phosphorus fractionation and sorption properties effected by drying pretreatments

3.1 Abstract

Stream sediment can control phosphorus (P) in the water column at baseflow. Two common laboratory analyses of sediment P are the equilibrium phosphate concentration at net zero sorption (EPC_0) and P fractionation. Good sample handling ensures representative results, but oftentimes studies rely on air-dried or freeze-dried samples, which alters sediment biogeochemistry. How and to what extent this influences EPC_0 and P fractionation remains unclear. We therefore examine pretreatment effects on sediment EPC_0 and P fractionation. We collected fine sediments (<2 mm) from streams in the Tukituki River and Reporoa basins in New Zealand (n = 31 sediments). Subsamples were then either kept fresh, frozen then lyophilized (freeze-dried), or dried at 40° C for two weeks (air-dried). Measurements of EPC_0 and P fractionation were made in triplicate. The sequential P fractionation scheme determined five different P pools: NH_4Cl (labile P); NaOH reactive P (RP; metal oxide-bound P) and unreactive P (URP; organic P); HCl (Ca-mineral P); and residual P. Along with statistical comparisons between fresh results and the two pre-treatments, we explored correlations between pre-treatment effects and sediment physicochemical characteristics. The sediments had generally low EPC_0 (majority <0.020 mg P L^{-1}), and uncertainty in EPC_0 increased with concentration magnitude. While there were sediment-specific changes in EPC_0 with pre-treatment, there was no consistent bias caused by pre-treatment. However, the differences between the fresh and air-dried sediment EPC_0 were larger and more variable than between fresh and freeze-dried sediment.

For P fractionation, the Tukituki sediments were enriched in HCl-P while Reporoa sediments had more NaOH-RP and NaOH-URP. Despite large sediment-specific changes, the overall effects of freeze- and air-drying sediment were: increased NH_4Cl-P (estimated average effect, $\hat{\theta}$ = +0.63 and +3.7 mg P kg^{-1}), no significant changes for NaOH-RP, contrasting changes in NaOH-URP (-3.4 and +3.3 mg P kg^{-1}), and decreased HCl-P (-40 and -33 mg P kg^{-1}). We found that drying sediment significantly influenced EPC_0 and P fractions (especially the NH_4Cl-P fraction). Air-drying was particularly error-prone and should be avoided. The use of freeze-drying to preserve samples for later analyses and improve ease of handling may be used with appropriate consideration of the research objectives and the error introduced by freeze-drying. However, we recommend using fresh sediments for analyses whenever possible, as they best represent natural conditions.

3.2 Introduction

Phosphorus (P) is a key limiting nutrient of primary production in aquatic ecosystems (Elser et al. 2007). Due to its numerous potential sources, variable chemical forms, and reactive transport in the environment, P pollution is difficult to target and mitigate (Sharpley et al. 2013; Powers et al. 2016). This transport is particularly complex in lotic systems, since numerous abiotic and biotic mechanisms control P fluxes (Reddy et al. 1999; House 2003; Withers and Jarvie 2008). Among abiotic factors, stream sediments are a major control of dissolved P in many streams (Jarvie et al. 2012; McDowell 2015). Since most fine sediments (<2 mm) have a great capacity for P adsorption (Barrow 1983a; Froelich 1988), and are themselves vectors for P derived from the original soil source (Condron and Newman 2011), studies of sediment-P interactions are important for characterizing the transport of P in streams and to receiving water-bodies.

Two common laboratory measurements for describing sediment-P interactions in streams are the equilibrium phosphate concentration at net zero sorption (Taylor and Kunishi 1971) and sediment P fractionation (Condron and Newman 2011; Wang et al. 2013). Equilibrium phosphate concentration at net zero sorption (EPC_0) is the estimate of the equilibrium dissolved reactive P (DRP) concentration in the solution of batch experiments containing sediment where neither net desorption nor adsorption occurs (Froelich 1988). The EPC_0 has been used to indicate the likely contribution of bed sediments in controlling water column DRP concentrations. For example, Jarvie et al. (2005) determined sediments to be an active sink for dissolved P at several stream sites subjected to wastewater treatment plant (WWTP) discharge since stream DRP concentrations were usually greater than the sediment EPC_0. Similar studies of streams with high P loading find enriched EPC_0 as an indication of previous P sequestration that might be released into solution again once in-stream DRP concentrations are reduced below the EPC_0 (Haggard et al. 2005; Ekka et al. 2006).

Sediment-P fractionation defines sequentially extracted pools of decreasingly bioavailable P. While some procedures may target compound-specific P (Golterman 1996), most produce operationally-defined fractions (Wang et al. 2013). Phosphorus fractionation provides valuable information on the forms of P being transported by the sediment and their reactivity (and potential bioavailability). For example, in some lakes and reservoirs, Ca-P can be the largest pool in sediments (identified by acid extractions), where the release to solution is mostly mediated by bacteria (Tang et al. 2014; Li et al. 2016) (Tang et al. 2014; Li et al. 2016). Fractionation can also provide important information on P reactions with sediments; Lin et al. (2009) found that P adsorbed to river sediments in isotherm experiments was primarily associated with Fe oxide minerals and (to a lesser extent) Al oxide minerals extracted by NaOH and NH_4F, respectively.

Data generated from either sediment EPC_0 or P fractionation relies on robust laboratory methods to be representative of the study system and comparable between studies. For EPC_0, previous work has highlighted that solution ionic strength and Ca^{2+} concentrations need to be similar to the study stream,

since low ionic strength and low Ca^{2+} (e.g., deionized water) can reduce the sediment's affinity for P (Klotz 1988; Rietra et al. 2001; Lucci et al. 2010). Additionally, pH is pivotal to the sorption process (Barrow 1983a), where lower pH generally increases adsorption affinity (Meyer 1979; Klotz 1988; Huang et al. 2016). A less significant factor is the temperature during incubation, where greater temperatures can increase reaction rates and EPC_0 (Barrow 1983a; Klotz 1988; McDowell et al. 2017). For P fractionation, the main variables lie with the choice of the fractionation scheme, where the investigator must consider: compound-specific versus operationally-defined P pools (Golterman 2002); the number of pools to measure; and the appropriate analytical methods to accurately measure concentrations (He and Honeycutt 2005; do Nascimento et al. 2016). Condron and Newman (2011) and Wang et al. (2013) provide helpful reviews on available P fractionation methods and guidance for choosing the appropriate method.

One laboratory variable that has not been formally addressed is the pre-treatment of sediment for storage and handling purposes. Often, fresh (wet) sediment is preferred for sediment-P analyses as changes that occur through drying are avoided (Haggard et al. 2007; Lottig and Stanley 2007; Condron and Newman 2011). However, logistical constraints often prevent the timely analysis or handling (e.g. weighing) of fresh sediment, so drying may be needed for long-term storage. For example, the intensive study design of stream P by Stutter et al. (2010) did not allow for all sediments to be analyzed fresh in the same timeframe; therefore, air-dried (30° C) sediments were used with the acceptance of the error introduced by drying. It is usually, but not always, acknowledged that air-drying sediment can alter redox properties (Phillips and Lovley 1987; Baldwin 1996), organic matter structure (Turner et al. 2007), and microbial content (Qiu and McComb 1995; Worsfold et al. 2005) of the sediment, and therefore is likely to alter P sorption and fractionation results (Klotz 1988; Condron and Newman 2011).

An alternative pre-treatment for sediments that is utilized by some protocols (e.g., Ruttenberg 1992) is freeze-drying. Although recommended by some workers for sediment-P analyses (Pettersson et al. 1988), freeze-drying is known to affect redox conditions (Phillips and Lovley 1987), some pools of nitrogen (Worsfold et al. 2008), and disrupt soil organic matter (Bartlett and James 1980). Some studies of P fractionation in lake sediments and similar systems suggest systematic differences between fresh and freeze-dried sediments (Barbanti et al. 1994; Goedkoop and Pettersson 2000), but little work has focused on the implications for stream sediments. Freeze-drying may provide a reasonable alternative for sediment storage when analysis of fresh sediments is not feasible.

This study discusses the effect of sediment pre-treatment on EPC_0 and P fractionation results. Using multiple stream sediments from two contrasting catchments, we compare results when analyzing sediments fresh, freeze-dried, and air-dried. We hypothesized that air-drying would produce the largest differences in EPC_0 and P fractions compared to fresh sediment data, but freeze-dried sediments would be more comparable to fresh sediments. Additionally, we hypothesized that freeze-drying and air-drying would produce results with less variation among replicate analyses than fresh samples.

3.3 Materials and methods

3.3.1 Study sites

Sediment sampling sites were located in two catchments on the North Island of New Zealand: the Tukituki and the Reporoa basin. The Tukituki basin is dominated (in New Zealand Soil Classification) by Brown and Pallic soils which equate to Dystrochrepts and Aquepts or Fragiochrepts, respectively, in US Soil Taxonomy (Hewitt 2010). Land use is dominated by high production exotic grasslands (77% as of 2012). The Tukituki basin receives approximately 800 mm of rainfall annually and mean annual temperature is 14.5°C. The Reporoa basin is dominated by Pumice soils (Vitrands in USDA soil taxonomy; Hewitt, 2010); land use (as of 2012) is predominantly dairy farming (44%) and exotic forestry (39%). Mean annual rainfall in the Reporoa basin varies from 1100 to 1550 mm and mean annual temperature is 12.6°C (Piper 2005).

Sampling took place during baseflow conditions in austral summer 2016/2017. Sediments were sampled from the submerged streambed at a variety of stream locations within each catchment (n=28) and from some floodplain sites in the Tukituki catchment (n=3). The Tukituki stream substrates were mostly gravel, sand, and some silt, while Reporoa stream substrates were sandy with a few cases of high silt (~80%); floodplain samples were predominantly sand. Further details about the study sites and their water quality can be found for the Tukituki catchment in (Quinn et al. 2020) and for the Reporoa catchment in McDowell et al. (2019b).

3.3.2 Sampling and preparation

Surficial sediments (uppermost 1 to 3 cm) were collected with a shovel during baseflow conditions. These sediments were located within the stream, near the centroid of flow, so as to target sediments under active flow. The stream sites, while at baseflow, were shallow and slow enough so as to prevent excessive winnowing of fine sediments during removal. We consider the streambed sediments to be oxic at time of collection as judged by the dissolved oxygen in the water-column (saturation was generally ≥100%) and sufficient streamflow (and thus hyporheic exchange). However, it is possible that anoxic micro-zones may be present, even at these shallow depths, depending on the biogeochemical context (Falco et al. 2016; Reeder et al. 2018). Samples were wet-sieved in the field to <2 mm with minimal exposure to air, kept cool (4° C) and in the dark during transit to the laboratory.

Approximately 5 g dry weight (d.w.) from each sediment sample was dried at 104° C to determine moisture content. Approximately 20 g d.w. of sediment was used for each pre-treatment. For the freeze-dried sediments, the subsample was first frozen (-20° C) for at least 24 h before being quickly transported to a freeze-drier for desiccation. The air-dried sediments were prepared by drying at 40° C (with ventilation) for two weeks. Air and freeze-dried samples were stored at 4° C until analysis.

For each pre-treatment, pH was measured in a 1:5 sediment to solution ratio (e.g., 1 g d.w. in 5 mL of solution) in D.I. water. For other physicochemical characteristics of the sediment, subsamples of the freeze-dried sediment were: microwave digested with nitric acid plus hydrogen peroxide, then analyzed via inductively coupled plasma optical emission spectrometry for total Al, Ca, Fe, Mg, K, Mn, Na, and Zn content; analyzed for total C content via a CN elemental analyzer; and analyzed for anion storage capacity (ASC; Blakemore et al. 1987).

All sediments (n = 31) were analyzed for P fractions, but only the Tukituki sediments (n = 20) were used for sorption experiments.

3.3.3 Sorption experiments

Batch experiments for EPC_0 were carried out in triplicate for each sediment following the methodology of Lucci et al. (2010). Solutions of KH_2PO_4 at 0, 0.01, 0.1, 1, 5, and 25 mg P L^{-1} were prepared with a background of 0.003 M $CaCl_2$ to simulate in-stream ionic strength and Ca^{2+} concentrations (Klotz 1988; Lucci et al. 2010). A pH buffer was not used for these solutions; however, all solutions were in the range 4.5-5.6. A higher pH (closer to the stream pH) would possibly alter the measured EPC_0 depending on the specific sediment surface chemistry regulating P uptake (Bolan and Barrow 1984; Klotz 1988; Huang et al. 2016). Using a sediment to solution ratio of 1:20, sediment samples were incubated with each solution for approximately 24 h via an end-over-end shaker near room-temperature (~20° C). After centrifugation (2400 *g* for 20 min), the supernatant was analyzed colorimetrically either via the molybdenum-blue method (Murphy and Riley 1962) or the malachite-green method (Ohno and Zibilske 1991; D'Angelo et al. 2001). We preferred the malachite-green method with a micro-plate reader for low concentrations (e.g., <0.1 mg P L^{-1}), since its detection limit (0.006 mg P L^{-1}; D'Angelo et al. 2001) is more sensitive than what we could achieve with the molybdenum-blue method on our spectrophotometer (~0.02 mg P L^{-1} for a 1 cm light-path). Standards prepared with the same background matrix were used for each batch of analyses. Adsorption or desorption of P was calculated as the mass of P either released into solution or removed from solution divided by the mass of sediment.

3.3.4 Phosphorus fractionation

We determined P fractions in each sediment using the Hieltjes and Lijklema (1980) method with some modifications. The sediments were sequentially extracted (at 1:100 sediment to solution ratio) with NH_4Cl (loosely sorbed P), NaOH (inorganic P associated with metal oxides; Danen-Louwerse et al. 1993), and HCl (inorganic P associated with calcium minerals). A wash step with NH_4Cl was included between NaOH and HCl extractions to prevent any significant carry-over (Condron and Newman 2011). The NaOH extract was also analyzed for total P (NaOH-TP; Pettersson et al. 1988) after an acid-persulfate autoclave digestion (USEPA 1978). After the HCl extraction, the residual pellet was dried, ground via a mortar and pestle, and digested for remaining P (residual P) by block digestion with H_2SO_4 and H_2O_2 (Olsen and Sommers 1982). The NaOH extracts were analyzed for reactive P (NaOH-RP)

with a modified molybdenum-blue method suitable for alkaline extracts (Dick and Tabatabai 1977; He and Honeycutt 2005), which prevents the hydrolysis of organic P known to occur with the analysis of reactive P via the single-solution molybdenum-blue method of Murphy and Riley (1962). Here, we denote the difference between NaOH-TP and NaOH-RP as unreactive P (NaOH-URP), which we consider to be primarily comprised of organic P (He and Honeycutt 2005). All other extracts and digested extracts were neutralized and then analyzed with the molybdenum-blue method of Murphy and Riley (1962).

Phosphorus fractionations were done in triplicate. Two separate observations of HCl-P (from two different sediment samples, one from the fresh and one from the freeze-dried pre-treatments) were censored due to extreme values (~800-1200 mg P kg^{-1} greater than other replicates and samples). Laboratory replicates were summarized with a geometric mean (to account for skewness) and used for statistical comparisons (see below).

3.3.5 Determination of EPC_0

Sediment EPC_0 was calculated as the *x*-intercept of the linear sorption model of P sorption on initial (rather than equilibrium) solution DRP concentration using only the solution concentrations from the 0 mg P L^{-1} treatment up to the lowest treatment with all triplicate points indicating positive sorption (in this study, either 0.01 or 0.1 mg P L^{-1}). Further explanation of the rationale behind this approach can be found in Appendix B.

In addition to calculating the EPC_0, the uncertainty about the EPC_0 was calculated in order to compare the variations induced by sediment pre-treatment. A 95% confidence interval about the *x*-intercept was estimated with a likelihood-based approach (see example 4 in Harding 1986). An example of this calculation is shown in Figure B.1.

3.3.6 Statistical comparisons

To test our hypothesis that sediment pre-treatment would affect P fractions and EPC_0, comparisons were made using the Wilcoxon signed-rank test (Hollander et al. 2013a). The null hypothesis of the test is that there is zero shift in location due to the pre-treatment (i.e., the paired differences have a distribution symmetric about a common median (θ) that is equal to zero); the data were treated as repeated measures where the T^+ statistic and *p*-value were computed for each comparison (fresh vs freeze-dried, fresh vs air-dried, and freeze-dried vs air-dried). Additionally, an estimate of θ ($\hat{\theta}$) and its 95% confidence interval were computed for evaluating the alternative case where $\theta \neq 0$ (Hollander et al. 2013a; Ugarte et al. 2015). The statistic $\hat{\theta}$ (which carries the same units as the observations) estimates, on average, both the magnitude and direction of pre-treatment effect when the null hypothesis is rejected.

Further, we explored the differences in P fractions and EPC_0 between pre-treatments with the sediment physicochemical data (given in Table C1, Appendix C). We used Spearman's ρ to measure the correlations between variables.

All analyses were performed in R (R Core Team 2020). The data and code used in this manuscript are provided online at https://doi.org/10.6084/m9.figshare.6157772.v1.

3.4 Results

3.4.1 EPC_0 and its uncertainty

The EPC_0 of the Tukituki sediments varied from 0.011 to 0.055 mg P L^{-1} for fresh sediments, 0.012 to 0.052 mg P L^{-1} for freeze-dried sediments, and 0.003 to 0.130 mg P L^{-1} for air-dried sediments (Figure 3.1). Most of these sediments had relatively low EPC_0 (<0.020 mg P L^{-1}) regardless of the pre-treatment used. The uncertainty about EPC_0 (here, the width of the 95% confidence interval for EPC_0) was correlated with the magnitude of EPC_0 (Spearman ρ = 0.439; Figure 3.2) where uncertainty was as low as 0.002 mg P L^{-1} and as high as 0.053 mg P L^{-1}.

It can be seen from Figure 3.1 that the disparity between pre-treatment methods varied for each sediment. For example, sediment T11 had practically identical results for each pre-treatment, while air-drying

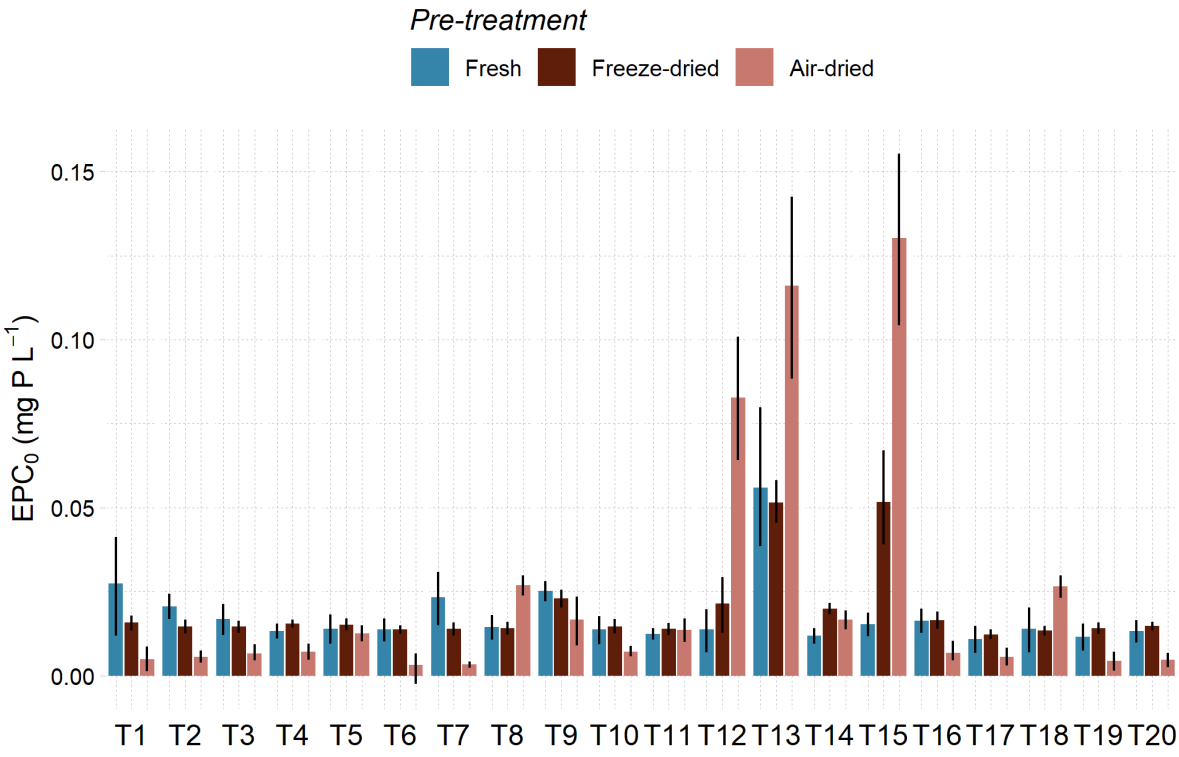

Figure 3.1 EPC_0 and the 95% confidence intervals estimated for EPC_0 for each pre-treatment (sediment sample ID is below each subplot).

reduced EPC_0 in other sediments such as T9 and T1. Additionally, the air-dried pre-treatment had the greatest uncertainty for sediment T9 while the fresh pre-treatment was the most variable for T1. Because effects of pre-treatment were different among each sediment, a Wilcoxon signed-rank test showed no *uniform* effect of pre-treatment for any of the three comparisons (p > 0.5; Table 3.1). This result is further illustrated in Figure 3.3.

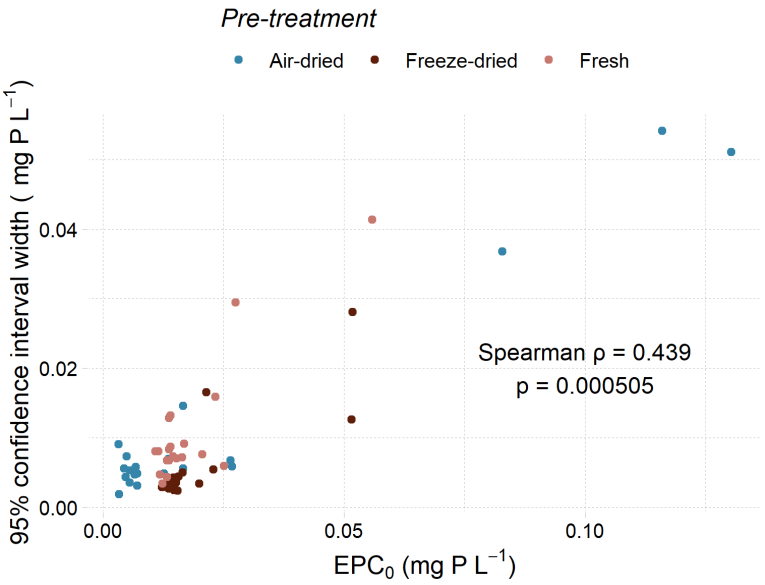

Figure 3.2 Uncertainty about EPC_0, as the width of the 95% confidence interval, compared to the magnitude of EPC_0; the Spearman rank correlation between EPC_0 uncertainty and EPC_0 is also given.

While differences in measured EPC_0 with magnitudes of approximately 0.010 mg P L^{-1} or more are ecologically relevant for sediment-P studies, the distribution of these differences is largely centered near zero ($\hat{\theta}$ of 0.0004 and -0.0032 mg P L^{-1} for the freeze-dried and air-dried comparisons, respectively; Table 1). Thus, no consistent bias can be attributed to pre-treatment based on our data.

However, we note that the variation in pre-treatment differences is considerably greater for air-drying in comparison to freeze-drying. In Figure 3.3, the standard deviation of comparisons and the inter-quartile range (IQR; difference between the 25th and 75th percentiles) are 0.010 and 0.004 mg P L^{-1} for the freeze-dried comparison; for the air-dried comparison, the standard deviation and IQR were 0.034 and 0.016 mg P L^{-1}. This result suggests that air-drying induces more deviation in EPC_0 from the fresh sediment EPC_0 than freeze-drying.

Tables C.1 and C.2 give the physicochemical characteristics of the sediments and their Spearman rank correlations with the EPC_0 values, respectively. Greater EPC_0 in the fresh sediments coincided with greater sediment Al, K, Mg, and Zn (at α=0.05). Changes in EPC_0 due to freeze-drying (positive values

Table 3.1 Wilcoxon signed-rank tests for the effect of sediment pre-treatment on sediment EPC_0 (n = 20). T^+ is the test statistic (i.e., sum of the positive ranks) used to calculate the p-value, where the null hypothesis is that the distribution of differences (second minus the first, e.g., freeze-dried EPC_0 minus fresh EPC_0) is centered about zero (θ=0); $\hat{\theta}$ is the pseudo-median calculated for the comparisons as an estimate for pre-treatment effect; the 95% confidence interval (C.I.) about $\hat{\theta}$ is also given.

Comparison	T^+	***p*-value**	$\hat{\theta}$ **(mg P L^{-1})**	**95% C.I.**	
Fresh vs. Freeze-dried	113	0.7841	0.000387	-0.00223	0.00208
Fresh vs. Air-dried	88	0.5459	-0.00315	-0.00860	0.02458
Freeze-dried vs Air-dried	90	0.5958	-0.0049	-0.00841	0.0258

being an increase in EPC_0 relative to the fresh sediment EPC_0) were negatively correlated with Al but positively correlated with anion storage capacity (ASC). For the air-dried pre-treatment, changes in EPC_0 were negatively correlated with sediment Al, Fe, Mg, and Mn content. There were no significant correlations of the EPC_0 measures with pH nor with changes in pH caused by drying.

3.4.2 Phosphorus fractionation

Sediment P fractions for each pre-treatment are shown for the Tukituki and Reporoa samples (Figure 3.4). The fresh Tukituki sediments were relatively low in the labile fractions, NH_4Cl and NaOH-RP (medians of ~0.4 mg P kg^{-1} and 19 mg P kg^{-1}, respectively). Unreactive P in the NaOH fraction was also low for Tukituki sediments (maximum of 19 mg P kg^{-1} but median of 4.4 mg P kg^{-1}). However, the Tukituki sediments were enriched in the HCl fraction (range of 250 to 410 mg P kg^{-1}). The fresh Reporoa sediments were also low in NH_4Cl-P (majority <0.5 mg P kg^{-1}), but had a relatively large amount of NaOH-RP (median of 280 mg P kg^{-1}). Additionally, the Reporoa sediments had medians of 9 mg P kg^{-1} in NaOH-URP and 140 mg P kg^{-1} in HCl-P.

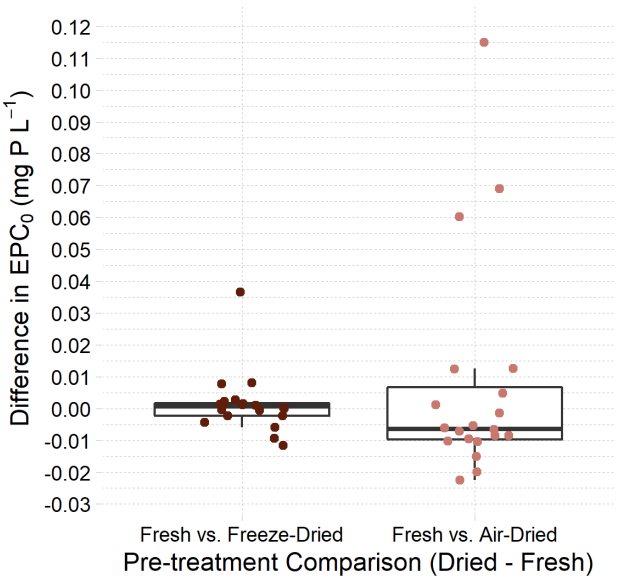

Figure 3.3 Boxplots of the differences in EPC_0 due to pre-treatment (dried minus fresh EPC_0; the underlying data is superimposed with arbitrary scatter for presentation purposes)

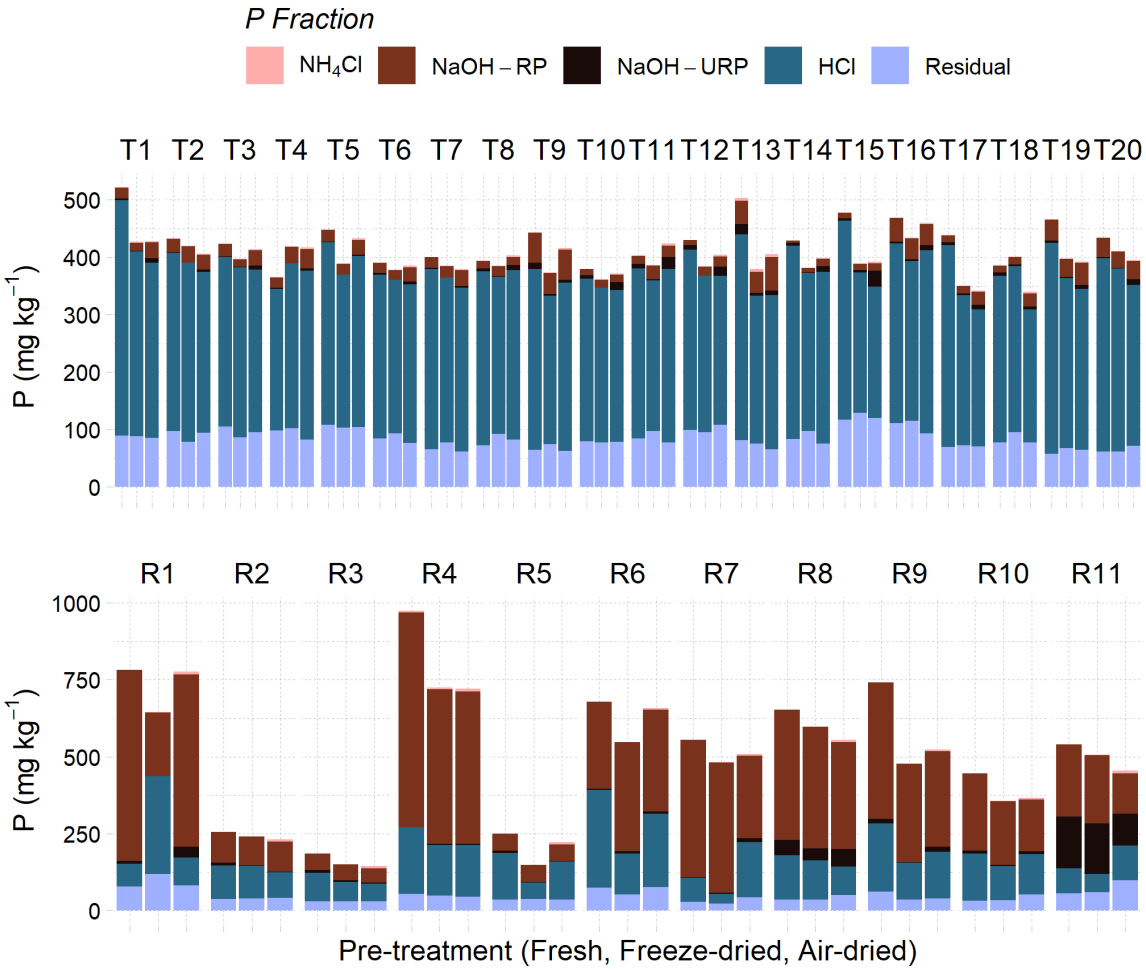

Figure 3.4 Phosphorus fractionations in stream sediments from the Tukituki basin (top row) and Reporoa basin (bottom row) for each pre-treatment (fresh, freeze-dried, and air-dried); note that within each sediment, the pre-treatments are ordered as fresh, freeze-dried, then air-dried and note the difference in scales for each row.

Sediment pre-treatment effects on P fractionation are summarized in Table 3.2. Both drying methods increased the NH_4Cl P fraction as compared to the fresh sediment, with an average increase of 0.63 mg P kg^{-1} and 3.7 mg P kg^{-1} for freeze-drying and air-drying, respectively. Further, the increase in NH_4Cl-P due to pre-treatment was 3.1 mg P kg^{-1} greater for air-drying than freeze-drying. Pre-treatment effects on the NaOH-RP fraction were mixed: changes due to either freeze-drying or air-drying were not significant at $\alpha = 0.05$ although Figure 3.4 suggests some decreases with drying – particularly for the Reporoa sediments (median values of differences with fresh Reporoa data were -22 and -76 mg P kg^{-1} for freeze-dried and air-dried, respectively). Results of NaOH-URP indicated a decrease with freeze-drying ($\hat{\theta}$=-3.4 mg P kg^{-1}), but an increase with air-drying ($\hat{\theta}$=3.3 mg P kg^{-1}). Both drying methods caused significant decreases in HCl-P in the sediments ($\hat{\theta}$=-40.4 and -33.3 mg P kg^{-1} for freeze-drying and air-drying, respectively). The residual P fraction was unaffected by drying methods.

Table 3.2 Wilcoxon signed-rank tests for effect of pre-treatment on each P fraction using all data (n = 31). T^+ is the test statistic (i.e., sum of the positive ranks) used to calculate the p-value where the null hypothesis is that the distribution of differences (second minus the first, e.g., freeze-dried sediment P minus fresh) is centered about zero (θ=0); $\hat{\theta}$ is the pseudo-median calculated for the comparisons as an estimate for pre-treatment effect; the 95% confidence interval (C.I.) about $\hat{\theta}$ is also given

P fraction	Pre-treatment Comparison	T+	p-value	$\hat{\theta}$ (mg P kg^{-1})	95% C.I.	
NH_4Cl	Fresh vs Freeze-dried	495	<0.001	0.6293	0.5043	0.7888
	Fresh vs Air-dried	496	<0.001	3.723	2.236	4.275
	Freeze-dried vs Air-dried	496	<0.001	3.134	1.548	3.602
NaOH-RP	Fresh vs Freeze-dried	149	0.0527	-4.107	-13.07	0.08472
	Fresh vs Air-dried	252	0.9461	0.5146	-37.88	5.028
	Freeze-dried vs Air-dried	282	0.5168	1.895	-5.804	6.018
NaOH-URP	Fresh vs Freeze-dried	50	<0.001	-3.433	-5.139	-1.867
	Fresh vs Air-dried	368	0.01766	3.266	0.7702	5.402
	Freeze-dried vs Air-dried	460	<0.001	5.734	3.692	9.242
HCl	Fresh vs Freeze-dried	50	<0.001	-40.43	-58.52	-24.6
	Fresh vs Air-dried	70	<0.001	-33.33	-51.39	-18.61
	Freeze-dried vs Air-dried	286	0.4678	4.909	-9.436	20.12
Residual	Fresh vs Freeze-dried	294	0.3777	2.128	-2.055	6.467
	Fresh vs Air-dried	235	0.8092	-0.3977	-4.107	3.664
	Freeze-dried vs Air-dried	209	0.4559	-1.941	-8.703	3.865

Spearman rank correlations of fresh sediment P and differences in sediment P due to pre-treatment with sediment characteristics for each P fraction are given in Table C.3 and are only summarized here. NH_4Cl-P content was positively correlated with metals (Al, Fe, Mg, Mn), but negatively correlated with ASC. While no correlations were apparent for freeze-drying, changes in NH_4Cl-P due to air-drying were negatively correlated with metals (e.g., greater amounts of P in the air-dried pre-treatment relative to fresh sediment corresponded with less amounts of Fe), but positively correlated with total C and ASC. NaOH-RP was negatively correlated with Ca, Mg, and pH, but positively correlated with Na, total C, and ASC; changes due to either drying method were negatively correlated with Na, ASC, and – just for the freeze-dried pre-treatment – total C. While some correlations were evident for the NaOH-URP, HCl-P, and residual P fractions in the fresh sediment, no noteworthy correlations were evident for the pre-treatment effects.

3.5 Discussion

3.5.1 Effects of pre-treatment on EPC_0

In this study, we report considerable variability in sediment EPC_0 dependent on the pre-treatment used (Figure 3.1, Figure 3.3), yet no consistent bias due to pre-treatment (Table 3.1). To our knowledge, few studies have examined changes in EPC_0 with drying. Klotz (1988) measured an increase in one sediment EPC_0 from 0.011 to 0.021 mg P L^{-1} with air-drying at 80°C – well within the differences calculated here (Figure 3.3). With a focus on lake sediments, Twinch (1987) compared fresh sediments and air-dried sediments (at room temperature) and measured significant increases in EPC_0 with drying (means of

0.095 and 0.284 mg P L^{-1}, respectively). These results cast doubt on the use of drying, particularly air-drying, for sediment P studies; however, the mechanisms for these changes in EPC_0 remain unclear.

For the case of an increased EPC_0 with drying (i.e., decrease of P affinity), at least two complementary biotic mechanisms are possible: lysing of microbial P (to be flushed when in contact with solution) and the removal of any potential biotic uptake by killing microbes. The flushing of microbial P upon drying has been demonstrated in lake sediments (Qiu and McComb 1995; Baldwin 1996) as well as after microbial death via less sediment-perturbing methods (e.g., autoclaving and irradiation; Meyer 1979; Klotz 1988; Haggard et al. 1999). McDowell (2003) showed an inverse relationship between microbial biomass P and desorbable P (as $CaCl_2$-extracted P) in stream sediments except where high organic carbon was present. Thus, the effects of both air-drying and freeze-drying could alter microbial P interactions and shift EPC_0 upward, particularly in sediments where biotic P uptake is relatively high (Lottig and Stanley 2007).

A possible abiotic mechanism for increasing EPC_0 with drying is the shift of sediment Fe from amorphous Fe-(oxyhydr)oxides to more crystalline Fe-oxides. Phillips and Lovley (1987) demonstrated that both air-drying and freeze-drying sediments oxidizes poorly-crystalline Fe (as measured by oxalate extracts). This oxidation could shift the Fe species from more amorphous, highly P-reactive Fe-(oxyhydr)oxides to more crystalline, less P-reactive Fe-oxides (Golterman 2004; Jan et al. 2015), i.e., the metal oxides are aged (see also discussion below). More crystalline metal-oxides may also explain the positive correlation observed between fresh sediment ASC (higher for more sorptive sediments) and increased EPC_0 after freeze-drying, although the same effect was insignificant for the air-dried pre-treatment Table C.2. Using lake sediments, Baldwin (1996) tested the effects of desiccation and oxidation (via air-drying) compared to oxidation alone on P adsorption and concluded that oxidation is the primary factor in reducing P affinity. The speciation of the sediment Fe in the current study was not measured, but average total Fe for the sediments in the EPC_0 experiments was 17.5 g kg^{-1} (standard deviation of 2.24 g kg^{-1}; n=20). Therefore, oxidizing and aging of sediment Fe may have increased EPC_0 in some of the dried pre-treatments, particularly the air-dried treatment (Figure 3.3).

Mechanisms for decreasing EPC_0 (i.e., an increase in P affinity) with drying are less clear. Sorption processes are typically stronger at lower pH (Barrow 1983a; Huang et al. 2016), which could be important for the cases of decreasing sediment pH with drying, but we found no relation between these variables for our data (Table C.2). As drying may affect organic matter content (Barbanti et al. 1994; Turner et al. 2007), an indirect effect may be tied to the competition between organic matter and phosphate for sorption sites (Guan et al. 2006).

Considering fresh sediment EPC_0 as the 'ideal' EPC_0, one objective was to identify a drying technique that induces the least variation in EPC_0 measurements. We have shown that freeze-drying introduces a modest amount of variation in EPC_0 (standard deviation of comparisons, 0.010 mg P L^{-1}; IQR of 0.004 mg P L^{-1}), but air-drying produces considerably more variation (standard deviation of comparisons,

0.034 mg P L^{-1}; IQR of 0.016 mg P L^{-1}; Figure 3.1). These figures are also reflected in the 95% confidence intervals for $\hat{\theta}$ in Table 3.1 (-0.0022 to 0.0021 and -0.0086 to 0.025 mg P L^{-1} for the freeze-dried and air-dried comparisons, respectively). Therefore, we recommend that, in cases where fresh sediment analyses are not practical, freeze-drying is the better alternative for sediment preservation before sorption analyses. We further recommend caution when using dried sediments for individual sorption studies where the goal is to characterize reactive transport processes (e.g., House and Denison 2002); the variability induced by drying may alter the interpretation of how the natural sediments behave *in situ*.

Additionally, as we have shown that uncertainty in EPC_0 increases with magnitude (Figure 3.2), care should be exercised in EPC_0 measurements not only for very low concentrations, but also for relatively high concentrations (e.g., > 0.030 mg P L^{-1}). Several studies examining impacted streams in a variety of settings have measured EPC_0 values ranging from near detection-limits (i.e., <0.010 mg P L^{-1}) to more than 1 mg P L^{-1} (Ekka et al. 2006; McDowell 2015; Weigelhofer 2017). Often, EPC_0 is used as an indicator variable, where changes in EPC0 are examined over multiple sampling periods or sites (Jarvie et al. 2005; Ekka et al. 2006), or used to correlate with baseflow DRP concentrations (McDowell 2015). These general uses of EPC_0 may not be significantly affected by the uncertainty we describe here, but other specific calculations may be more tenuous (e.g., percentage EPC saturation; Jarvie et al. 2005).

3.5.2 Effects of pre-treatment on P fractionation

Drying pre-treatments of stream sediments can alter P fractions thus confounding important P biogeochemical processes. We have demonstrated that, generally, the largest P fraction in a sediment is the most susceptible to error caused by drying (e.g., HCl-P in the Tukituki sediments and NaOH-RP in the Reporoa sediments; Figure 3.4). However, error in smaller P fractions (e.g., NH_4Cl-P) should not be ignored as these pools can represent highly bioavailable pools (Pettersson et al. 1988; Condron and Newman 2011; Wang et al. 2013). Indeed, the labile pools such as NH_4Cl-P are critical as they also correlate with P available for exchange with the water column (McDowell 2015). In this study, we showed consistent increases in NH_4Cl-P (0.63 to 3.7 mg P kg^{-1}; Table 3.2) with either drying pre-treatment, with air-drying being the most severe. This is likely due to either lysing of microbial P (Qiu and McComb 1995; McDowell 2003) or reduction in sediment P affinity (Baldwin 1996), but NH_4Cl is also capable of dissolving some small amounts of $CaCO_3$, Al, or Fe oxide-bound P (Hieltjes and Lijklema 1980; Pettersson et al. 1988) – there is not enough information to discern if any interactions between drying and labile calcite or metal P compounds took place.

Sediments with larger amounts of NaOH-RP (i.e., the Reporoa sediments; mean of 330 mg P kg^{-1}) had the largest changes in the NaOH-RP fraction; however, there was no consistent pre-treatment effect for all sediments (Table 3.2). The fact that more of the sediments studied here were relatively poor in NaOH-RP (i.e., the Tukituki sediments; mean of 22 mg P kg^{-1}) may have weakened any possible effects of pre-treatment on this fraction. Decreases in this somewhat labile fraction with either freeze-drying or air-

drying have been reported, where the P may instead appear in more labile fractions (Dieter et al. 2015) or potentially become occluded by either particle aggregation (Twinch 1987) or aging metal oxides (Schlichting and Leinweber 2002; Hjorth 2004), thus appearing as more recalcitrant P. While the fractionation scheme we employed here does not partition between various metal oxide-bound P fractions (e.g., amorphous Fe oxy(hydr)oxides extracted first with bicarbonate-dithionite (BD) followed by extraction of more crystalline Fe and Al oxides with NaOH; Jan et al. 2015), it is conceivable that oxidation and aging of metal oxides incurred by either drying method would alter the speciation and lability of metal oxide-bound P (Phillips and Lovley 1987). In fact, when studying air-drying effects on sediment P, Dieter et al. (2015) used a similar P fractionation scheme to that of the current study but included a BD step: both NH_4Cl-P and BD-P increased after the studied lake sediments were air-dried, concomitant with a decrease in NaOH-RP (reduced or redox-insensitive Fe and Al oxide bound P in this context) and NaOH-URP. Therefore, when drying sediments from a strongly reduced environment (many lakes, wetlands, and some streams; Reddy et al. 1999) or when using a more detailed P fractionation scheme, pre-treatment effects may be magnified.

The sediments studied here were relatively low in alkaline-extracted URP (overall mean of 12 mg P kg^{-1}), as compared to other New Zealand stream sediments (average NaOH-URP ranging from 35 to 57 mg P kg^{-1}; McDowell and Hill 2015). Yet, there was a moderate decrease in NaOH-URP with freeze-drying ($\hat{\theta}$ = -3.4 mg P kg^{-1}) and increase with air-drying ($\hat{\theta}$ = 3.3 mg P kg^{-1}; Table 3.2). Turner et al. (2007) studied the effects of air-drying and freeze-drying on NaOH-EDTA extracts for wetland soils and also had conflicting results: between pre-treatments, there were sample-specific changes in both total P recoveries and speciation (via ^{31}P nuclear magnetic resonance spectroscopy; e.g., changes in proportions of phosphate and various organic P compounds). It remains unclear what mechanisms are most important, as potential drying effects include enhancing organic P recoveries in alkaline extracts by increasing organic P lability or by disrupting organic matter as well as the potential to solubilize some organic P (particularly during air-drying) so that it is instead detected as part of the labile P fractions (Barbanti et al. 1994; Cade-Menun 2005; Turner et al. 2007; Dieter et al. 2015).

The HCl-P fraction largely contains various Ca-P minerals (e.g., apatite) and is typically stable (Wang et al. 2006), but small amounts may be mobilized for lower pH at the microscale (Golterman 2004). Both pre-treatments resulted in large decreases in this fraction (-40 and -33 mg P kg^{-1} for freeze-drying and air-drying, respectively). Schlichting and Leinweber (2002) reported similar findings for P fractionation in a peat soil: acid-extracted P was significantly reduced for freeze-drying and air-drying pre-treatments, which the authors attributed to decreased solubility of Ca-phosphates.

3.6 Conclusions

If EPC_0 is to be a useful parameter for describing sediment-P interactions in streams, then methods should be as robust and replicable as possible. While previous work has recommended using solutions that match the stream chemistry, we further recommend using fresh sediments for EPC_0 measurements

whenever possible. However, in cases of logistical and handling constraints, freeze-drying should be a preferred storage method (with tacit acknowledgement of potential errors of approximately 0.01 mg P L^{-1}). In agreement with our hypothesis, air-drying increased variability and uncertainty compared to fresh (and freeze-dried) sediments and should be avoided in the measurement of EPC_0.

For P fractionation, we recommend that the pre-treatment of samples should be uniform within a given study and that, when comparing results to studies employing differing pre-treatments, caution should be used. In particular, we have shown that labile P fractions (e.g., NH_4Cl-P) are the most susceptible to changes with drying, where air-drying had the most dramatic effect. Given the changes in P fractions that can occur with any drying method, we also share the past recommendations that fresh sediment data should be the most representative of *in situ* conditions and the favored pre-treatment.

Chapter 4

Phosphorus attenuation in streams by water-column geochemistry and benthic sediment reactive iron

4.1 Abstract

Streams can attenuate inputs of phosphorus (P) and, therefore the likelihood of ecosystem eutrophication. This attenuation is poorly understood in reference to the geochemical mechanisms involved. In our study, we measured P attenuation mechanisms in the form of (1) mineral (co-)precipitation from the water-column and (2) P sorption with benthic sediments. We hypothesized that both mechanisms would vary with catchment geology and, further, that P sorption would depend on reactive iron (Fe) content in sediments. We sampled 31 streams at baseflow conditions, covering a gradient of P inputs (via land use), hydrological characteristics, and catchment geologies. Geochemical equilibria in the water-column were measured and benthic sediments were analyzed for sorption properties and P and Fe fractions. Neither P-containing minerals nor calcite-phosphate co-precipitation had the potential to form. However, in-stream dissolved reactive P (DRP) correlated with labile sediment P (water-soluble and easily-reduced Fe-P), but only for streams where hyporheic exchange between the water-column and sediment porewaters was likely sufficient. Because this labile P was associated with poorly-crystalline Fe oxides, which determined P sorption capacity, we observed that more sorptive sediments were positively, rather than negatively, related to DRP concentrations. Sediment labile P concentrations normalized to sorption capacities, however, were uncorrelated with DRP. Our results suggest that the combination of biogeochemical Fe and P cycles and the hydrological exchange with the hyporheic zone attenuate DRP in these streams at baseflow. Such combinations will likely vary spatiotemporally within a catchment and must be considered alongside inputs of P and sediment if the P concentrations at baseflow – and eutrophication risk – are to be well managed.

4.2 Introduction

Once mobilized from land via surface runoff or sub-surface flows (McDowell et al. 2004), phosphorus (P) is repeatedly impeded along its flowpath by biotic and abiotic processes (Haggard and Sharpley 2007; Baulch et al. 2013). This persistence of P gives rise to 'legacy P' (Sharpley et al. 2013; Powers et al. 2016), where past P inputs can take unknown years (decades, centuries) to deplete thereby masking the effects of mitigation efforts (Meals et al. 2010; Crockford et al. 2015).

At baseflow, P in streams is subject to biotic and abiotic processes which could lead to the transient storage of P (e.g., associated with sediments or stored in biomass) and its re-mobilization back to the water column (Dodds 2003; House 2003). Both periphyton (Biggs 2000; Dodds and Smith 2016) and heterotrophic microbes (Mulholland et al. 1997; McDowell 2003) can assimilate P, especially dissolved reactive P (DRP; mostly orthophosphate but can include labile organic compounds), as it is the most bioavailable P form (Biggs 2000; Muscarella et al. 2014). Numerous studies on biotic P attenuation, however, conclude that biotic mechanisms may not comprise all P attenuation in many streams (Hall et al. 2002; Weigelhofer et al. 2018a; Griffiths and Johnson 2018). This highlights the contribution of abiotic (i.e., geochemical) mechanisms for P attenuation (Jarvie et al. 2006a; McDaniel et al. 2009; Stutter et al. 2010). Yet, abiotic P attenuation in streams is poorly studied. To address this gap, we consider two major geochemical mechanisms for P attenuation: calcium (Ca) based (co-)precipitation and sediment P sorption.

Calcium-phosphate mineral precipitation and $CaCO_3$ co-precipitation may remove DRP from the water column given sufficient Ca, pH, and *p*CO2 (Stumm and Morgan 1996; House 2003; Golterman 2004). Like carbonates and other minerals formed when the water-column is super-saturated in respect to those phases, these minerals could remove P from the stream, with the precipitants mostly depositing on benthic or macrophyte surfaces in the stream (Golterman 2004; Parker et al. 2007; Corman et al. 2015). This may contribute to the initial removal of DRP from the water column, and further adsorption or mineral transformations (e.g., towards hydroxyapatite) may occur (Golterman 1988; Diaz et al. 1994; Plant and House 2002). Given the dependence on water-column geochemistry, Ca-based P attenuation is likely a function of catchment geology (Mulholland et al. 1997; House 2003; Corman et al. 2015). Indeed, most studies focused on Ca-based P attenuation are located in catchments with strongly calcareous (i.e., chalk, karst) geology (Diaz et al. 1994; House 1999; Jarvie et al. 2006a; Cohen et al. 2013). Few studies have considered a range of geologies where other abiotic P attenuation mechanisms may prevail, particularly sediment P sorption.

Sorption onto benthic stream sediments is likely a common component of P attenuation (Froelich 1988; Haggard and Sharpley 2007; McDowell 2015), especially for baseflow conditions where water is given time to contact sediments in the hyporheic zone (Harvey 2016). Sorption intensity often correlates negatively with in-stream DRP (McDaniel et al. 2009; McDowell 2015; Weigelhofer et al. 2018a) and positively with stream P uptake metrics (Jarvie et al. 2005; Haggard et al. 2005; Demars 2008).

Likewise, streams with high P loading (i.e., high sustained DRP concentrations) tend to have sediments with diminished sorption capacity and greater stores of P (Jarvie et al. 2012; McDowell 2015), particularly in the more labile and redox-sensitive pools (Lewandowski and Nützmann 2010; Weigelhofer 2017).

These redox-sensitive pools (i.e., iron (Fe) oxy(hydr)oxides, henceforth Fe oxides) are pivotal for P attenuation in streams, where dynamic redox fronts in the hyporheic zone encourage cycling of Fe oxides (Boano et al. 2014; Peiffer et al. 2021). Along with clay minerals and other metal (e.g., aluminum) oxides, Fe oxides are among the reactive surfaces responsible for P sorption (Parfitt 1979; Golterman 2004; Gérard 2016) but may dominate P sorption sites in sediments of many non-calcareous streams (Lewandowski and Nützmann 2010; van der Grift et al. 2014; Dupas et al. 2018b). Amorphous, surface-reactive Fe oxides – e.g., poorly-crystalline goethite, lepidocrocite, and ferrihydrite (Stumm and Sulzberger 1992; Jan et al. 2015) – have the greatest affinity for P (Lijklema 1980; Goldberg and Sposito 1984b). Yet, these Fe oxides usually compose only a minority of total sediment Fe (Hyacinthe et al. 2006; Jan et al. 2013). Much stream sediment research focuses on total Fe or roughly defined Fe fractions to discuss Fe-P interactions, thus limiting our understanding of how Fe oxides affect DRP concentrations (Hoffman et al. 2009; Rawlins 2011; Tye et al. 2016; Kreiling et al. 2019). In other environments, amorphous Fe oxides are increasingly recognized as predominant P sorption sites (Marton and Roberts 2014; Parsons et al. 2017). We hypothesize this to apply to stream benthic sediments as well.

In the present study, we examined P attenuation mechanisms in streams at baseflow via Ca-P geochemistry in the water column, stores of sediment P and redox-sensitive Fe, and P sorption capacities of sediments. Given that these processes are likely all tied to catchment geology and P inputs, we sampled waters and benthic sediments of streams from a variety of geologies, land use, and stream characteristics and which differ in typical baseflow DRP concentrations. We hypothesized that the primary abiotic mechanisms responsible for P attenuation (and therefore related to DRP concentrations) were Ca-based mineral equilibria in the water column and sorption with benthic sediments. Under this hypothesis, we expected that the prominence of either mechanism would be tied to geology: Ca-P (co-)precipitation would be more likely for streams draining calcareous geologies and sediments would be more sorptive for P when originating from geologies rich in Fe and Al minerals. Further, we hypothesized that amorphous, reactive Fe determined sediment P sorption, rather than refractory or total Fe pools. Specifically, our first objective was to identify whether Ca-P (co-)precipitants were favorable in streams draining calcareous geologies (here, sedimentary geologies) and, hence, a potential mechanism for P attenuation. Our second objective was to identify the major pools of P in the sediments in relation to their potential lability (e.g., sensitivity to redox) and how they vary between catchments. Our third objective was to model sediment P sorption as a function of Fe fractions to test whether this mechanism for P attenuation varied with the reactivity of the Fe.

4.3 Materials and Methods

4.3.1 Study sites

We sampled 31 streams in the Canterbury region, New Zealand (Figure 4.1). We targeted sites to cover a variety of catchment characteristics and typical baseflow DRP concentrations (McDowell et al. 2013). The site characteristics are given in Table S1, according to the River Environment Classification (REC) developed for New Zealand (Snelder and Biggs 2002). Stream sizes were mostly 1st to 5th order, with some 6th and 7th order streams (n=7), and generally included both low-elevation and hilly/mountainous streams. Most of the catchments contained some amount of pastoral land uses (i.e., sheep and dairy farming), reflecting the dominant land use in the Canterbury region. In north Canterbury, basins are characterized by quaternary fluvial gravels with some underlying sedimentary deposits (e.g., limestone); further south, the plains were formed by river-deposited erosion and glacial outwash products (by the Rakaia and Waimakariri Rivers), with some intermittent outcrops of greywacke; Banks Peninsula is characterized by a basalt volcanic geology (Brown 2001). Catchment geology of the study streams corresponded to alluvium (n=15), sedimentary (hard and soft sedimentary; n=10), and volcanic basic (n=6). Further, we used the REC to distinguish between two prominent sources of flow for the streams by identifying spring-fed sites (n = 8) from the other sites (termed here as 'hill-fed'; n = 23).

4.3.2 Sampling

Sediment and stream water samples were collected at each site during baseflow conditions from March to May, 2018 (austral late summer/early autumn). We sampled between 1000 and 1600 h to minimize possible diel effects on DRP (Cohen et al. 2013; McDowell et al. 2019b). Benthic sediments (top 1-3 cm) were collected *in situ* with a scoop and wet-sieved in the field to <2 mm. The sieved sediment slurry settled after 30 minutes, where excess water was decanted and ~2 kg of wet sediment was stored on ice and later refrigerated in the laboratory (4 °C). Sediment sampling locations were targeted within the stream where surface water interacts with benthic sediments; primarily, riffle beds near the centroid of flow were sampled but depositional areas closer to the bank were sampled at sites where the stream was too deep and fast-flowing for practical sampling.

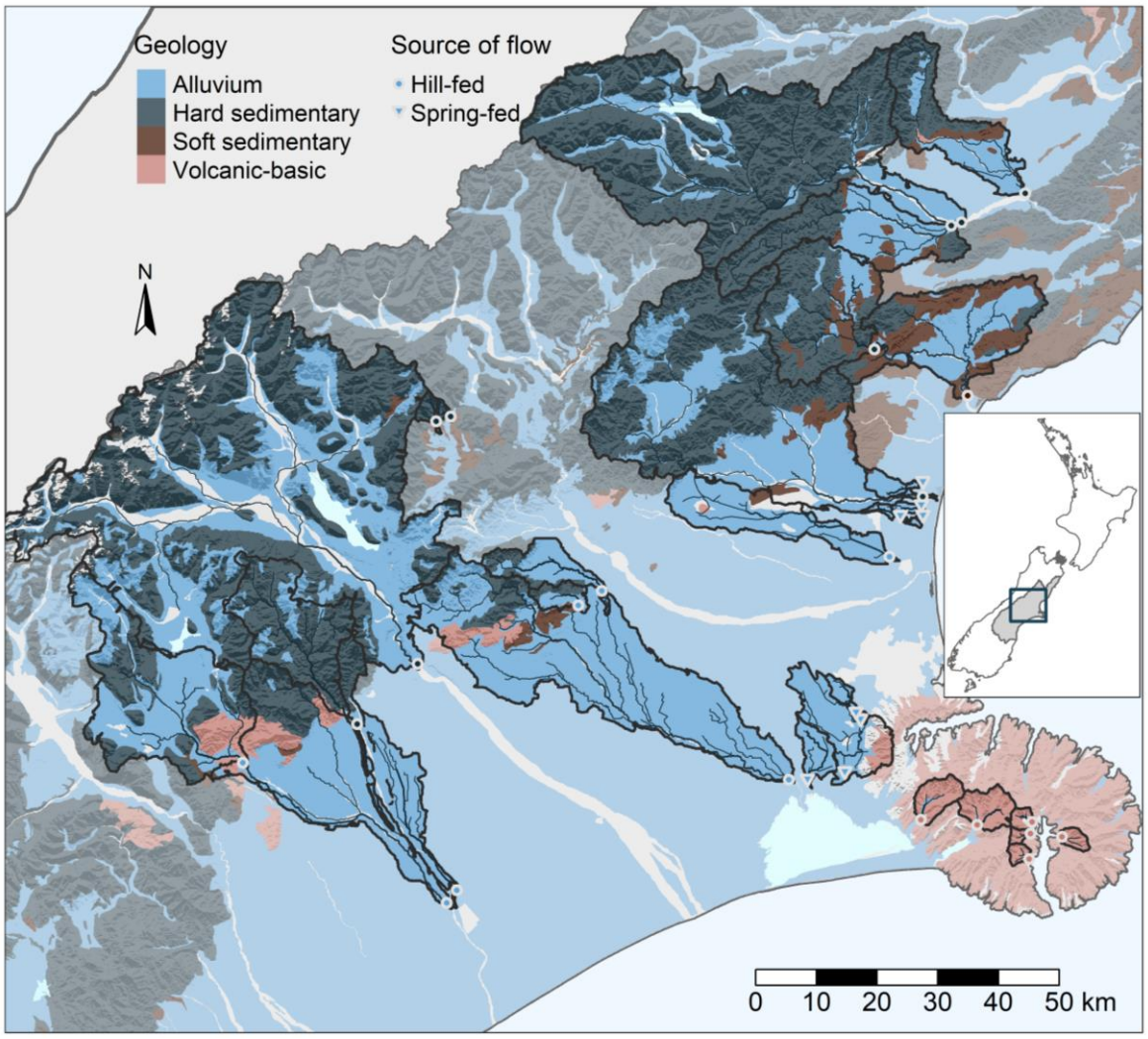

Figure 4.1 Study streams, and their catchments, in Canterbury, New Zealand (see inset). The geology classification used by the (New Zealand) River Environment Classification scheme is shown, with 'Miscellaneous' (loess, peat), 'Other' (river beds, ice cover, lakes, and urban centers), and Plutonic geology classes excluded for clarity. Sampling locations are shown, with circles indicating 'Hill-fed' streams (i.e., no major groundwater inputs) and triangles indicating 'spring-fed' streams.

For water grab samples, bottles were field-rinsed three times before taking a water sample at two-thirds of the stream depth. Two subsamples were filtered in the field (0.45 µm) while another subsample was left unfiltered (all with minimal headspace); all samples were then stored on ice for transport back to the laboratory followed by either freezing (-20 °C; for ion chromatography, ICP-OES, and dissolved Fe as explained below) or refrigeration (4 °C). In addition, we measured dissolved oxygen and temperature at each stream.

4.3.3 Water physicochemical analyses

Upon return to the laboratory, we immediately measured pH and conductivity in the unfiltered stream sample. Alkalinity was measured on the filtered sample within 24 h of collection (Rounds 2012). We measured DRP on the filtered sample within 24 h via the malachite-green method (detection limit of 0.006 mg P L^{-1}; Ohno and Zibilske, 1991; D'Angelo et al., 2001). We measured dissolved anions (F, Cl, SO_4, NO_3) via ion chromatography (detection limits range 0.02 to 0.50 mg L^{-1}), cations (Al, Ca, Fe, K, Mg, Mn, Na, Zn) via ICP-OES (detection limits approximately 0.002 mg L^{-1}), and total dissolved Fe via

a ferrozine method (see below); other elements (e.g., some trace metals) were below detection. Total suspended solids (APHA 2005) were low in these streams at baseflow (mean of 7 mg L^{-1}), so suspended sediment likely had negligible influence on DRP (data not shown). Blanks and duplicate checks were included in each batch of samples to ensure quality of results.

For total Fe in filtered water and sediment extracts (below), we modified the ferrozine colorimetric method (Stookey 1970; Viollier et al. 2000) as it provided greater sensitivity than ICP-OES for concentrations <0.050 mg Fe L^{-1} (Appendix D.1). The method detection limit was approximately 0.010 mg Fe L^{-1} in solutions and 0.017 mg Fe L^{-1} for digests.

4.3.4 Sediment physicochemical analyses

We refer to sediments kept wet as 'fresh'. A subsample was dried at 104 °C to measure moisture content. Fresh sediment pH was measured in D.I. water at 1:5 sediment to solution ratio. We measured two common P sorption indices on fresh sediment: anion storage capacity (ASC; Blakemore et al. 1987) and the Bache-Williams index (BWI; Bache and Williams 1971) as modified by Burkitt et al. (2002). Both are single-point isotherms with overnight shaking (16 h) but ASC is a measure of P retention at low pH while BWI is a measure of P retention at neutral pH and controlled Ca concentration.

A subsample of sediment was freeze-dried and analyzed for total C and N with a CN elemental analyzer and for total element concentrations with ICP-OES following microwave digestion with nitric acid plus hydrogen peroxide. Additionally, dried sediments were sieved to <1 mm and analyzed for particle size distributions via laser-diffraction, which are expressed on a percent volume basis (Eshel et al. 2004); here, we define clays as ≤ 4 μm, silt as > 4 μm and ≤ 62.5 μm, and sands as > 62.5 μm.

4.3.5 Sediment phosphorus and iron fractionation

We began sediment P fractionation within a week of sampling using fresh sediments (Simpson et al. 2019). We followed the scheme of Jan et al. (2015) as it distinguishes between amorphous, reactive Fe oxides and more crystalline or recalcitrant Fe phases but we also included a fraction from the SEDEX scheme (Ruttenberg 1992) as described below. The sequential fractionation is summarized in Table 4.1; see further details in Appendix D.1. We used 0.5 g (dry weight) sediment and 10 mL of extractant in each step, in triplicate. The bicarbonate-dithionite (BD) solution (0.1 M $NaHCO_3$ and 0.1 M $Na_2S_2O_4$, pH 7.2; BD-I and BD-II fractions) was prepared fresh with degassed D.I. water (subject to vacuum for 30 min) and used immediately. The NaOH (I and II) and HCl fractions used 1 M NaOH and 0.5 M HCl, respectively. A 0.5 M NaCl wash step was included after the BD-II and NaOH-II steps to prevent carryover to the next fraction (Condron and Newman 2011). The BD-I and NaOH-I fractions involved 5 min of shaking, immediate centrifuging, decanting, and a further 5 min extraction (i.e., 10 minute total

Table 4.1 Experimental procedure of the sequential sediment P fractionation; an additional fraction estimated via a complementary scheme (SEDEX) is also shown. The targeted biogeochemical pools of P are given but are not exact since fractionation methods are operationally defined. P and Fe analyses in the BD and NaOH fractions refer to total P and Fe. †Not applicable

Step or fraction	Solution	Extraction Time	Analyses	Primary biogeochemical pool
Jan et al. (2015) scheme:				
1. H_2O	D.I. water	30 min	P	Labile, loosely-bound P
2. BD-I	Bicarbonate-dithionite	5 + 5 min	P, Fe	Less-crystalline, surface-active Fe oxides
3. BD-II	Bicarbonate-dithionite	2 h	P, Fe	Crystalline, poorly active Fe oxides
Wash	NaCl	5 min	NA†	NA
4. NaOH-I	NaOH	5 + 5 min	P, Fe	Active Al oxides, labile organic matter, clay minerals
5. NaOH-II	NaOH	16 h	P, Fe	Crystalline Al oxides, refractory organic matter, clay minerals
Wash	NaCl	5 min	NA	NA
6. HCl	HCl	24 h	P	Primary minerals
Modified SEDEX (Jensen et al 1998) scheme:				
Acetate	Acetate buffer (pH=4)	6 h	P	Authigenic apatite and $CaCO_3$-bound P, leachable organic matter

extraction time) as per Jan et al. (2015). Here, we consider the sum of H_2O-P and BD-I P to be labile sediment P. To complement our dataset regarding Ca-P phases formed *in situ*, we included the acetate buffer step (pH of 4) from the SEDEX scheme as modified by Jensen et al. (1998). Following the extraction at room temperature with an end-over-end shaker, we centrifuged (10 min at 2400 *g*) and filtered the extracts (Whatman grade 41).

Total P in the BD and NaOH fractions was determined via acid-persulfate autoclave digestion (method 4500-P; APHA 2005) followed by the molybdenum-blue method (method detection limit of ~0.02 mg P L^{-1} for digests). Total Fe in the BD and NaOH digests was measured with the ferrozine method described above. We examined patterns in Fe and P contents among the BD and NaOH fractions with molar Fe:P ratios. All fractionation data presented are averages of laboratory triplicate analyses.

4.3.6 Geochemical equilibria

We employed the PHREEQC geochemical software (Parkhurst and Appelo 2013) with the MINTEQA2 version 4 database to calculate mineral saturation indices (SIs) and therefore discuss saturation states with reference to the minerals (Appelo and Postma 2005). A mineral's SI is defined as $\log_{10}(IAP/K_{sp})$, where IAP is the ion activity product measured in solution and K_{sp} is the mineral's equilibrium solubility constant. A SI >0 and SI <0 indicate super- and subsaturation with respect to the mineral phase. We employ these data to detect if the thermodynamic equilibria favor precipitation reactions as a potential mechanism for DRP removal, but cannot determine what phases actually occur as there may be kinetic limitations (Stumm and Morgan 1996; Plant and House 2002). The analytical input data consisted of stream temperature, pH, total dissolved anions and cations, alkalinity, and DRP.

4.3.7 Data and statistical methods

Two sites in the sedimentary class – both at pristine, forested headwaters – had DRP concentrations below our detection limits. To mitigate potential bias, we inserted reference DRP values based on similar streams according to the REC (McDowell et al. 2013). Additionally, through exploratory analyses, we found it necessary to remove data with a spring-fed source of flow as a confounding variable when modelling DRP. We hypothesized that the spring-fed streams had greatly diminished hyporheic exchange, thus limiting the interaction between sediment reaction sites and the water column. This limitation could be due to 1) accumulation of fine, silty sediments which restricted hydraulic conductivity (Packman and Salehin 2003; Weigelhofer et al. 2018a), 2) low-gradients which limited hydrodynamic forces at the streambed (Boano et al. 2014), and 3) the likely presence of groundwater inputs along the reach, which are known to limit hyporheic exchange flows (Azizian et al. 2017).

We calculated a proxy variable similar to degree of P saturation (Pierzynski 2000), where sediment labile P (H_2O-P + BD-I P) was divided by ASC and then scaled for ease of interpretation. We lack detailed Al data and hence overlook Al-bound P in this metric, however, this normalized sorption saturation is likely comparable to the degree of P saturation.

We summarized differences in stream and sediment chemistry between the three geology classes with nonparametric tests (Hollander et al. 2013b). We first tested the null hypothesis (H_0) that the locations of the group-wise distributions were equal via the Kruskal-Wallis test; if H_0 was rejected, we then constructed multiple comparisons with rank statistics, adjusting for simultaneous inferences (Konietschke et al. 2015). Spearman correlations were used to describe relationships between variables of interest. We used robust linear models ("rlm" function in MASS package; Venables and Ripley 2002) to fit simple predictive models for DRP and sediment P sorption (see Appendix D.2). All tests were performed at 95% confidence and all analyses were conducted in R (version 3.5.2; R Core Team 2020). The data are available at Figshare (https://doi.org/10.6084/m9.figshare.11999922.v1).

4.4 Results

4.4.1 Stream chemistry and mineral equilibria

Stream waters ranged from low to moderate alkalinity (16.7 to 88.8 mg $CaCO_3$ L^{-1}, mean of 42.7 mg $CaCO_3$ L^{-1}) and mean conductivity was 143 µS cm^{-1} (Table 4.2). For the time of sampling (generally, 1000 to 1600 h), pH averaged 7.44 (6.43 to 7.93). Dissolved reactive P varied among geology classes with the lowest median DRP concentration in sedimentary, followed by alluvium, and then by volcanic basic.

Stream solution composition was examined by calculating geochemical equilibria via PHREEQC (Figure D1). Stable, yet kinetically slow-to-form, Al and Fe minerals such as gibbsite and goethite were supersaturated in the water column; amorphous Al species ($Al(OH)_3(am)$) were subsaturated. The most reactive Fe species studied here, ferrihydrite ($Fe(OH)_3(am)$), was supersaturated (overall mean saturation index (SI) 1.9 ± 0.6), particularly over volcanic basic geologies (mean SI 2.9 ± 0.2). We note that ferrihydrite SI correlated with log-activity of HPO_4^{2-} ($\rho = 0.71$, $p < 0.001$). Further, $CaCO_3$ precipitation was not favorable (SI range of -2.1 to -0.08). Therefore, for these grab samples, $CaCO_3$ co-precipitation of phosphate was not a likely mechanism.

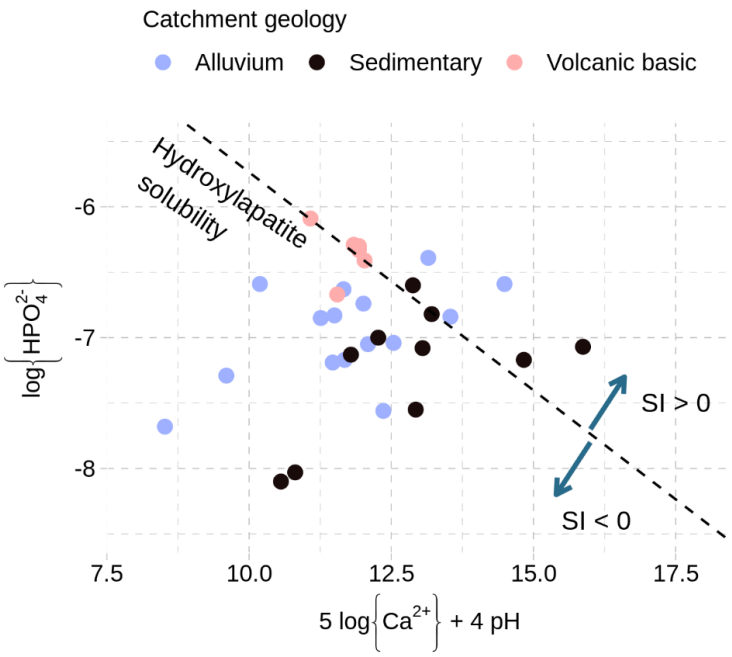

Figure 4.2 Log-activity of HPO_4^{2-} plotted against a function of log-activities of Ca^{2+} and H^+ for the stream samples (n=31). The dashed line is a reference solubility line for hydroxylapatite, where points below (above) this line are likely sub-saturated (super-saturated) with respect to hydroxylapatite, as indicated by the saturation index (SI).

The activities of phosphate and Fe/Al were too low to precipitate any P minerals from the water column in these streams (e.g., strengite; Figure D.1). For Ca-phosphate minerals, the less thermodynamically stable phases (e.g., $CaHPO_4 \cdot (H_2O)_2$) were sub-saturated in all samples. However, several streams

Table 4.2 Summary of stream water chemistry grouped by River Environment Classification geology class; values are given as medians (means ± standard deviation); lowercase letter exponents represent group-wise comparisons between geology classes.

		Geology Class		
	Units	Alluvium (n=15)	Sedimentary (n=10)	Volcanic basic (n=6)
pH	S.U.	7.29^a (7.26 ± 0.33)	7.54^b (7.61 ± 0.22)	7.62^b (7.62 ± 0.06)
Conductivity	µS cm^{-1}	142 (154 ± 51)	78 (121 ± 79)	155 (152 ± 18)
Alkalinity	mg L^{-1} as $CaCO_3$	42.1 (44.3 ± 11.6)	31.4 (42.9 ± 23.8)	39.4 (38.2 ± 6.8)
DRP	µg L^{-1}	7.4^b (10.8 ± 7.6)	5.0^a (5.6 ± 2.8)	27.7^c (26.9 ± 11.3)
NO_3-N	mg L^{-1}	1.947^b (1.857 ± 1.22)	0.111^a (0.626 ± 1.18)	0.152^a (0.151 ± 0.080)
SO_4	mg L^{-1}	8.21^b (8.44 ± 2.77)	4.89^b (9.38 ± 7.44)	3.54^a (3.49 ± 0.55)
Dissolved Fe	mg L^{-1}	0.011^a (0.06 ± 0.139)	0.011^a (0.012 ± 0.006)	0.137^b (0.153 ± 0.069)
Dissolved Ca	mg L^{-1}	18.1^b (18.3 ± 5.52)	11.9^b (17.09 ± 9.8)	8.9^a (8.79 ± 1.33)

showed saturation to slight super-saturation with respect to hydroxylapatite (Figure 4.2). All three geologies had some waters with positive hydroxylapatite SIs but, notably, the saturation appears to not extend significantly past the hydroxylapatite solubility curve (max SI of 1.7). We noted no other geochemically significant relationships for other mineral SIs (Figure D.2) or ion log-activities (Figure D.3). Given that most phosphorous minerals of interest were sub-saturated in the water column, and that hydroxylapatite likely requires greater SI values before actively precipitating (see discussion), mineral precipitation does not seem a significant mechanism for P removal in these streams.

4.4.2 Stream sediment physicochemistry

The benthic sediments were largely neutral (mean pH of 7.10; Table 4.3). The fine sediments sampled in this survey were predominantly sandy (mean sand content of 84%), although five alluvium sites and two volcanic basic sites had less than 80% sand content (silt + clay content range from 21 to 92%). However, sediments from spring-fed streams were much finer (D_{50} of 171 ± 172 µm) than those in hill-fed streams (597 ± 172 µm). Similarly, sediments were relatively low in total C (overall median of 3.0 g C kg^{-1}) except for the spring-fed sites (median of 20 g C kg^{-1}). Owing to the different geological origins, the volcanic basic sediments were enriched with Al, Fe, Mn, and P in comparison to the sedimentary and alluvium sediments.

Table 4.3 Select physicochemical properties of the stream benthic sediments grouped by River Environment Classification geology class; D_{50} is the median particle size of the fine sediments (<2 mm); values are given as medians (means ± standard deviation); lower case letter exponents represent group-wise comparisons between geology classes.

		Geology Class		
	Units	Alluvium (*n*=15)	Sedimentary (*n*=10)	Volcanic basic (*n*=6)
pH	S.U.	6.93^{a} (6.91 ± 0.24)	7.12^{ab} (7.26 ± 0.46)	7.26^{b} (7.31 ± 0.26)
Total C	g kg^{-1}	3.68^{b} (13.2 ± 17.9)	1.06^{a} (1.3 ± 0.58)	8.53^{b} (10.8 ± 5.9)
Total N	g kg^{-1}	0.41^{b} (1.31 ± 1.69)	0.20^{a} (0.23 ± 0.12)	0.65^{b} (0.83 ± 0.43)
Total Al	g kg^{-1}	20.15^{a} (21.93 ± 6.04)	18.44^{a} (20.42 ± 4.88)	35.39^{b} (37.2 ± 8.35)
Total Ca	g kg^{-1}	5.77^{a} (6.25 ± 2.34)	5.97^{ab} (6.35 ± 2.21)	9.59^{b} (9.46 ± 2.14)
Total Fe	g kg^{-1}	22.14^{a} (21.95 ± 4.05)	23.31^{a} (23.44 ± 3.36)	48.45^{b} (50.43 ± 5.93)
Total Mn	mg kg^{-1}	379^{a} (455 ± 275)	372^{a} (368 ± 89.2)	725^{b} (713 ± 160)
Total P	mg kg^{-1}	548^{a} (597 ± 219)	476^{a} (462 ± 121)	2130^{b} (2220 ± 525)
Clay	% volume	0.011^{b} (0.99 ± 1.6)	0^{a} (0.001 ± 0.004)	0.3^{b} (0.5 ± 0.7)
Silt	% volume	5.25^{b} (19.8 ± 25.9)	0.049^{a} (0.98 ± 1.54)	14.7^{b} (16.8 ± 9.96)
Sand	% volume	94.7^{a} (79.2 ± 27.4)	99.9^{b} (99.0 ± 1.55)	85.0^{a} (82.7 ± 10.7)
D_{50}	µm	415^{a} (356 ± 255)	701^{b} (680 ± 232)	544^{ab} (491 ± 238)

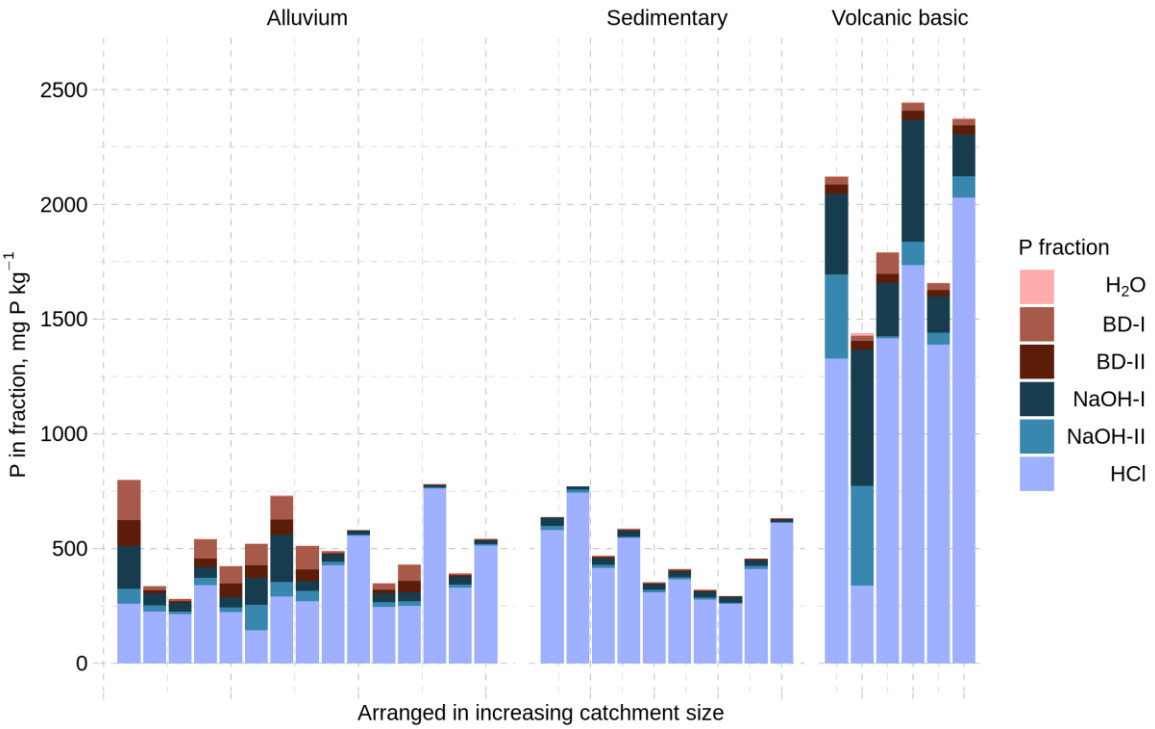

Figure 4.3 Benthic sediment (<2 mm) phosphorus content fractionated according to decreasing chemical lability (via the Jan et al. 2015 scheme); bars are arranged in increasing order of catchment size within the River Environment Classification geology class.

4.4.3 Sediment phosphorus and iron fractionation

Phosphorus fractionation

Sediment P fractionation varied considerably by the catchment geology (Figure 4.3). The most labile pool, H_2O-P, was relatively high in the volcanic basic sediments and alluvium sediments, but lower for sedimentary sites (Table 4.4). Both reductively-soluble sediment P pools, BD-I and BD-II, were enriched in the alluvium and volcanic basic sediments relative to sedimentary sites, with total BD-extractable P (BD-I plus BD-II) of 84.3 ± 84.1, 77.6 ± 26.7, and 6.60 ± 3.03 mg P kg^{-1}, respectively. Both BD-I and BD-II P fractions correlated with fines (clays plus silts) concentration (respectively, ρ = 0.89 and 0.84, $p < 0.001$ for both tests; Figure C.4). The labile P pool (H_2O-P and BD-I P together) composed only 0.5 to 21% of total P.

The NaOH-extractable P pool averaged 34.3 ± 10.7, 91.8 ± 84.7, and 518 ± 329 mg P kg^{-1} for sediments from the sedimentary, alluvium, and volcanic basic geologies, respectively. Total NaOH-extractable P correlated with Al content (ρ = 0.47, p = 0.008), but no other relationships (e.g., with total C) were evident. The least available sediment P pool analyzed, HCl-P, was greatest in volcanic basic sediments and correlated strongly with Ca (ρ = 0.84, $p < 0.001$) and Mg content (ρ = 0.71, $p < 0.001$).

Table 4.4 Stream sediment P fractions, Fe fractions, molar Fe:P ratios (including total Fe to total P), and sorption metrics. Values are given as medians (means ± standard deviation). The lowercase letter exponents represent group-wise comparisons between geology classes.

		Geology Class		
	Units	Alluvium (*n*=15)	Sedimentary (*n*=10)	Volcanic basic (*n*=6)
Sediment P fractions				
H_2O	mg P kg^{-1}	1.39^{b} (1.82 ± 1.41)	0.417^{a} (0.436 ± 0.081)	2.48^{b} (3.84 ± 3.76)
BD-I		27.8^{b} (52.1 ± 51.9)	4.14^{a} (3.62 ± 1.87)	31.8^{b} (40.4 ± 25.5)
BD-II		14.8^{b} (32.2 ± 32.6)	3.53^{a} (2.98 ± 1.405)	38.4^{b} (37.2 ± 4.52)
NaOH-I		39.4^{b} (61.2 ± 60.3)	25.2^{a} (23.8 ± 7.11)	292^{c} (342 ± 185)
NaOH-II		20.2^{ab} (30.6 ± 29.1)	10.5^{a} (10.4 ± 4.79)	97.2^{b} (176 ± 177)
Acetate		12.9 (22.9 ± 20.0)	12.8 (12.1 ± 4.05)	25.4 (22.0 ± 7.53)
HCl		271^{a} (337 ± 163)	414^{ab} (452 ± 161)	1400^{b} (1370 ± 572)
Sediment Fe fractions				
BD-I	mg Fe kg^{-1}	529^{b} (1360 ± 1500)	252^{a} (254 ± 66.1)	2480^{c} (2560 ± 623)
BD-II		664^{a} (915 ± 930)	433^{a} (446 ± 190)	5990^{b} (5150 ± 1670)
NaOH-I		50.4^{b} (92.98 ± 102)	21.7^{a} (22.6 ± 8.18)	107^{c} (118 ± 41.6)
NaOH-II		447^{b} (122 ± 182)	35.9^{a} (36.2 ± 11.1)	99.6^{b} (131 ± 94.6)
Fe:P (molar)				
BD-I	mol Fe mol P^{-1}	15.5^{a} (18.3 ± 9.85)	36.4^{b} (65.01 ± 66.9)	39.1^{b} (40.6 ± 14.3)
BD-II		19.9^{a} (31.6 ± 25.05)	80.7^{b} (99.2 ± 61.01)	85.8^{b} (76.6 ± 21.6)
NaOH-I		0.71^{a} (1.15 ± 1.15)	0.52^{b} (0.59 ± 0.34)	0.23^{b} (0.25 ± 0.15)
NaOH-II		2.11 (2.69 ± 2.40)	1.80 (2.47 ± 1.63)	0.38 (3.12 ± 6.25)
Total		22.3^{b} (22.2 ± 5.73)	28.4^{c} (29.1 ± 4.25)	13.3^{a} (13.1 ± 2.64)
Sorption metrics				
ASC	% P retention	14.3^{b} (33.6 ± 26.9)	6.98^{a} (10.2 ± 5.58)	49.4^{b} (49.4 ± 11.1)
BWI	$\frac{\text{mg P kg}^{-1}}{\log_{10}(\mu\text{g P L}^{-1})}$	12.7^{b} (51.95 ± 54.7)	6.84^{a} (8.8 ± 5.57)	53.5^{b} (53.4 ± 12.3)

In a separate analysis, we estimated an authigenic Ca-P fraction (acetate P) distinct from the HCl-P fraction. Acetate P did not differ between geologies (overall median of 13.1 mg P kg^{-1}). Acetate P correlated with finer sediments ($\rho = 0.62$, $p < 0.001$) and total C content ($\rho = 0.63$, $p < 0.001$) but not with DRP nor calcite SI.

For hill-fed streams, in-stream DRP correlated well with labile P (Figure 4.4). Using geology and sediment P pools as the primary predictors for DRP in the hill-fed streams (excluding P sorption potential, see below; Table C.2), the best-fit linear model employed geology and H_2O-P (RMSE = 3.93 µg P L^{-1}), with a slope (95% C.I.) of 2.62 (1.58 to 3.66) µg P L^{-1} per mg P kg^{-1} for H_2O-P. Competing

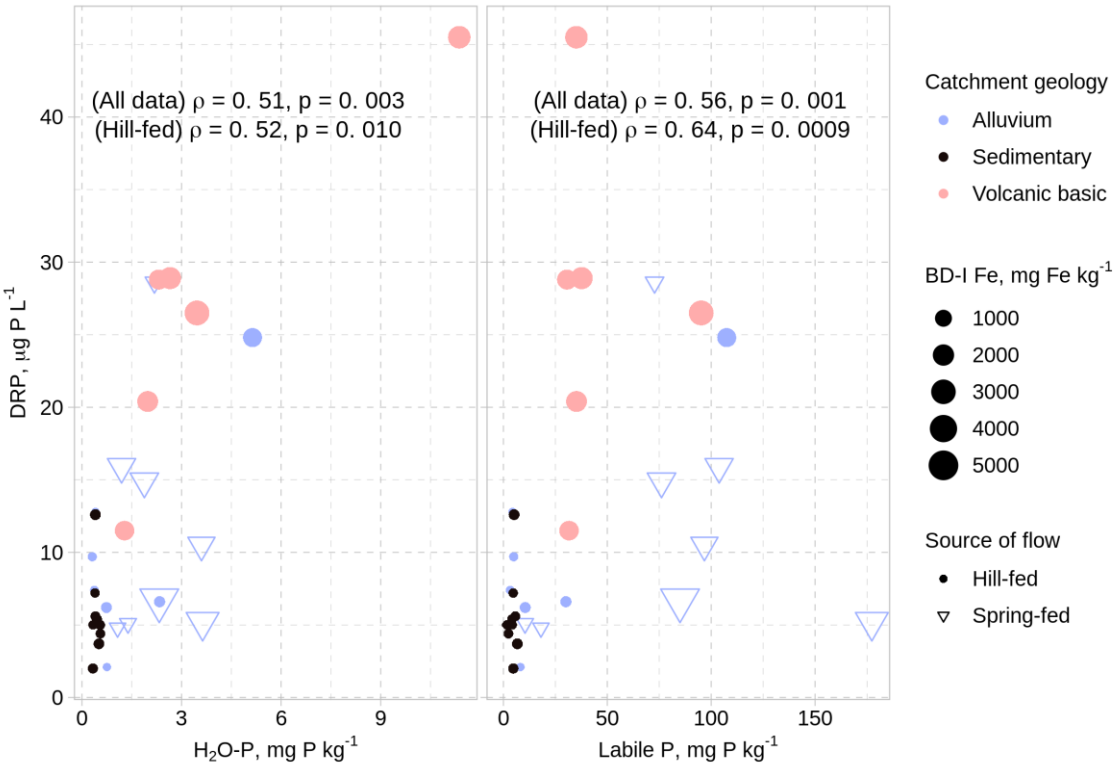

Figure 4.4 Stream DRP concentration as a function of sediment H_2O-P and sediment labile P (H_2O-P plus BD-I P) concentrations from the sequential P fractionations. Spearman correlation tests are shown for all data (*n*=31) as well as for only hill-fed streams (*n*=23).

models either had much greater AIC (models with either geology or H_2O-P only) or similar AIC but more model degrees of freedom (model with geology, H_2O-P, and BD-I P).

Iron fractionation

Although total Fe was similar between alluvium and sedimentary sites (Table 4.3), reactive Fe pools varied markedly within and between all catchment geologies (Table 4.4; Figure 4.5). Total BD-extractable Fe was 2280 (± 2200), 700 (± 242), and 7710 (± 2120) mg Fe kg^{-1} for the alluvium, sedimentary, and volcanic basic sites, respectively. On average, the amorphous Fe pool (BD-I) made up 52, 38, and 34% of the BD-extractable Fe in alluvium, sedimentary, and volcanic basic sediments. BD-extractable Fe correlated with percent fines (Figure D.5), where Spearman ρ was 0.87 ($p < 0.001$) and 0.79 ($p < 0.001$) for BD-I and BD-II, respectively. The Fe extracted by NaOH was at least one order of magnitude less than BD-extractable Fe for each geology (totals of 59 to 250 mg Fe kg^{-1}); NaOH-I Fe related to percent fines ($\rho = 0.77$, $p < 0.001$) and total C ($\rho = 0.75$, $p < 0.001$), which is likely due to release of Fe complexed with organic matter during extraction. Total Fe concentrations were correlated with both BD-extractable ($\rho = 0.45$; $p = 0.012$) and NaOH-extractable Fe ($\rho = 0.39$; $p = 0.033$). However, these reactive Fe pools contributed little to total Fe in these sediments (8.6 ± 7.2% and 0.6 ± 0.8% for total BD- and NaOH-extractable Fe, respectively; Figure 4.5).

Fe:P ratios in sediments

Molar Fe:P in pools of sediment P varied depending on the pool and the catchment geology (Table 4.4; Figure D6). In general, Fe:P ratios were much lower for the NaOH fractions (approximately 0.25 to 3 mol Fe mol P^{-1}) than for other fractions (>13 mol Fe mol P^{-1}), likely due to little surface reactive Fe left after BD extraction while containing P bound with constituents other than Fe. A paired Wilcoxon signed-rank test on Fe:P ratios in the BD fractions indicated that BD-II Fe:P is, on average, 25.6 (95% C.I., 14.4 to 36.1) mol Fe mol P^{-1} greater than BD-I Fe:P. However, Fe:P ratios for either BD-I or BD-II fractions showed little ability to predict either DRP or ASC (data not shown). For the BD (I and II) fractions, median Fe:P was less in alluvium sediments compared to both sedimentary and volcanic-basic sediments.

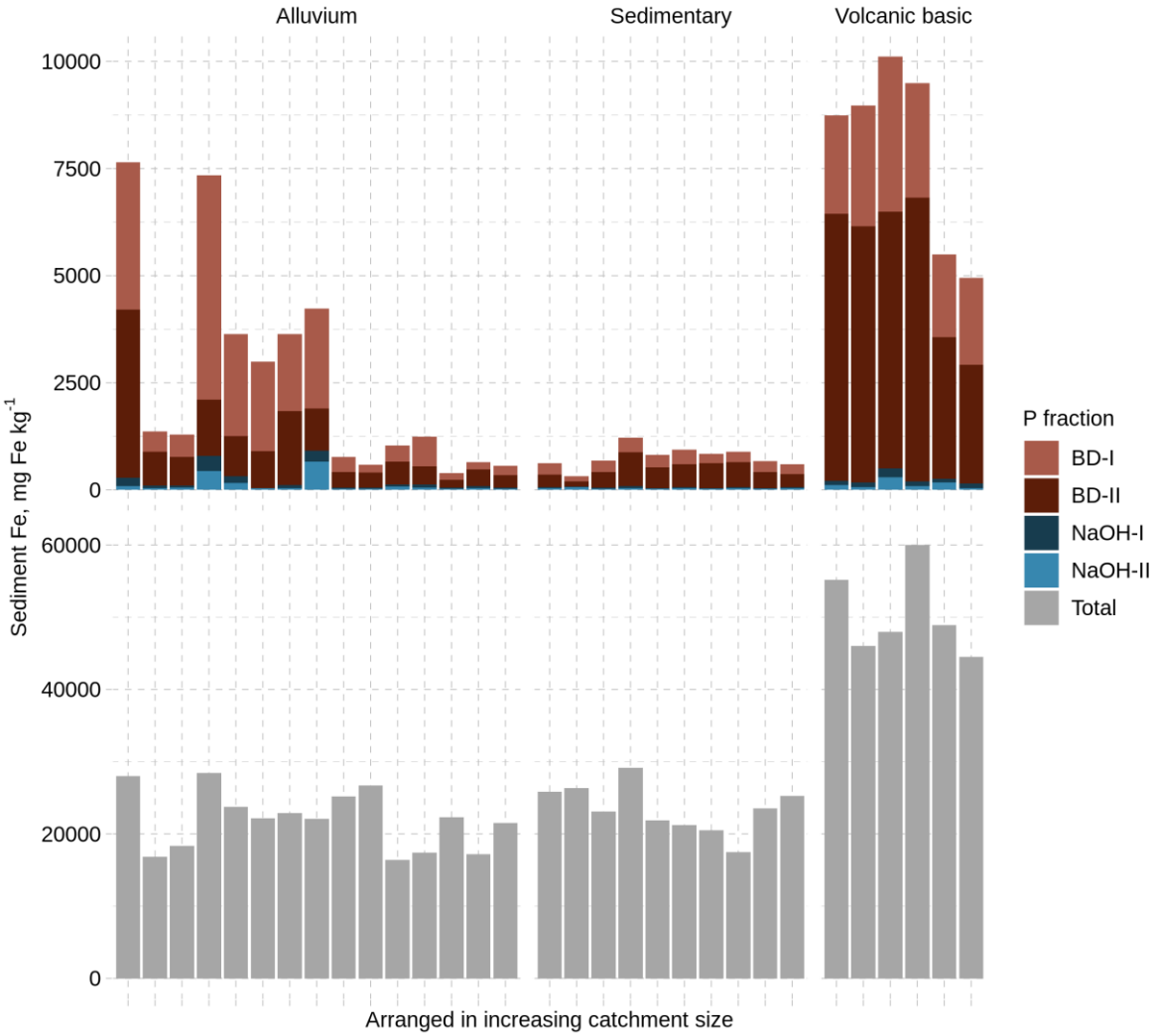

Figure 4.5 Sediment Fe concentration as measured in the P fractions (potentially reactive Fe arranged in decreasing lability; top row) or as the total Fe concentration (bottom row); note the differences in scale. Each bar (site) is arranged in increasing catchment size within the River Environment Classification geology class.

4.4.4 Stream sediment phosphorus sorption

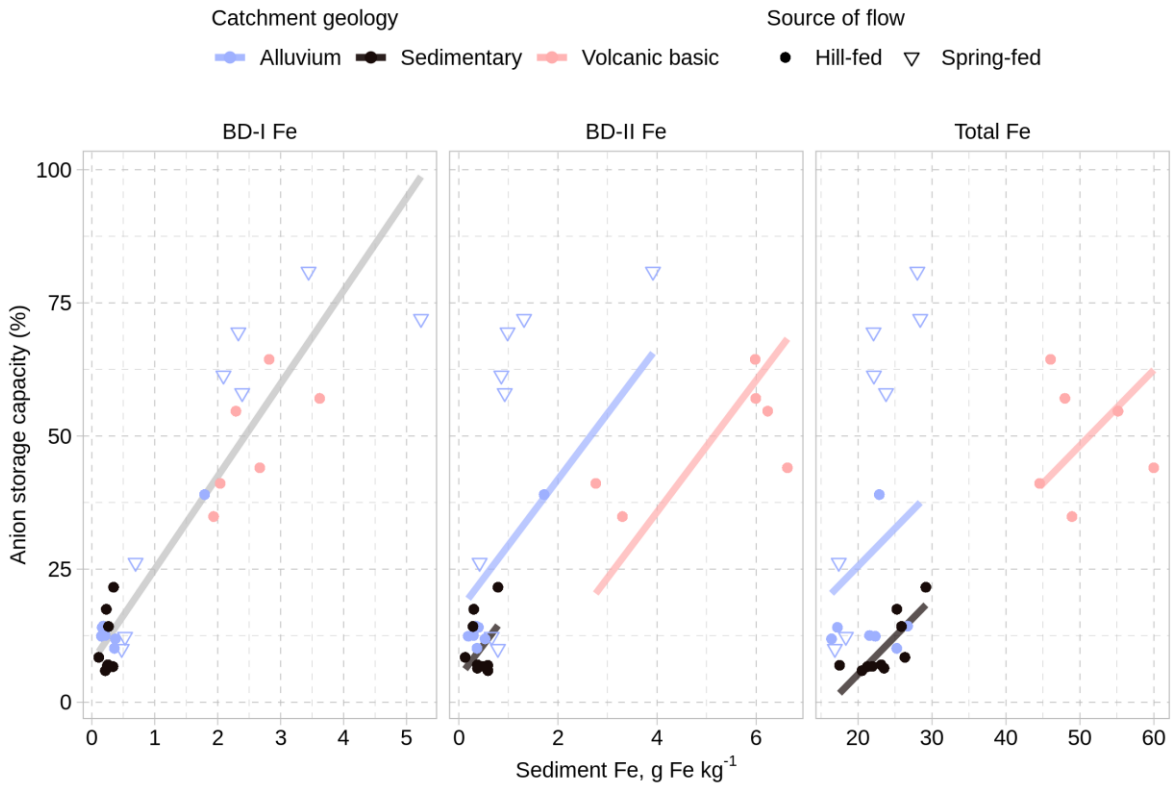

Figure 4.6 Sediment P sorption potential as measured by anion storage capacity (ASC; %) as a function of sediment Fe in the bicarbonate-dithionite (BD) extractable pools and total sediment Fe. The optimal robust linear models for ASC with each Fe fraction, as discussed in text, are shown; note that while BD-I Fe alone predicts ASC, the models for BD-II Fe and total Fe include catchment geology as a covariate.

Although ASC and BWI differ in how P sorption potential is determined, a similar pattern was apparent in both variables for the three sampled geologies; we focus on ASC for brevity. The sedimentary samples had lower ASC than either alluvium or volcanic basic sediments (Table 4.4). There was no clear relationship between ASC and sediment Fe:P ratios (Figure D.7), while ASC was correlated with BD-I ($\rho = 0.77$, $p < 0.001$), BD-II ($\rho = 0.65$, $p < 0.001$), and total Fe ($\rho = 0.55$, $p = 0.002$) pools (Figure 4.6).

More refractory pools of Fe were less predictive of ASC than BD-I Fe (Figure 4.6). While ASC for the refractory Fe pools tended to cluster according to geology and stream source of flow, ASC varied linearly as a function of BD-I Fe (all data), which we illustrate with a linear modelling exercise (Table C3). Among the simpler models (i.e., single covariates), the model of only BD-I Fe had the lowest AIC and the lowest RMSE (9.2 %); the estimated slope for BD-I Fe was 0.0174 (0.0152 to 0.0197) % per mg Fe kg^{-1}. While adding catchment geology did not improve the model fit for BD-I Fe, it did improve the fit for models involving either BD-II or total Fe. Thus, BD-I Fe alone predicted ASC well and captured the variance otherwise provided by geology or refractory Fe to the lesser models. Adding acetate-P as a proxy for sediment carbonate phases (see discussion) to the model with BD-I Fe slightly improved the fit by lowering AIC (231 to 228) and RMSE (8.4%).

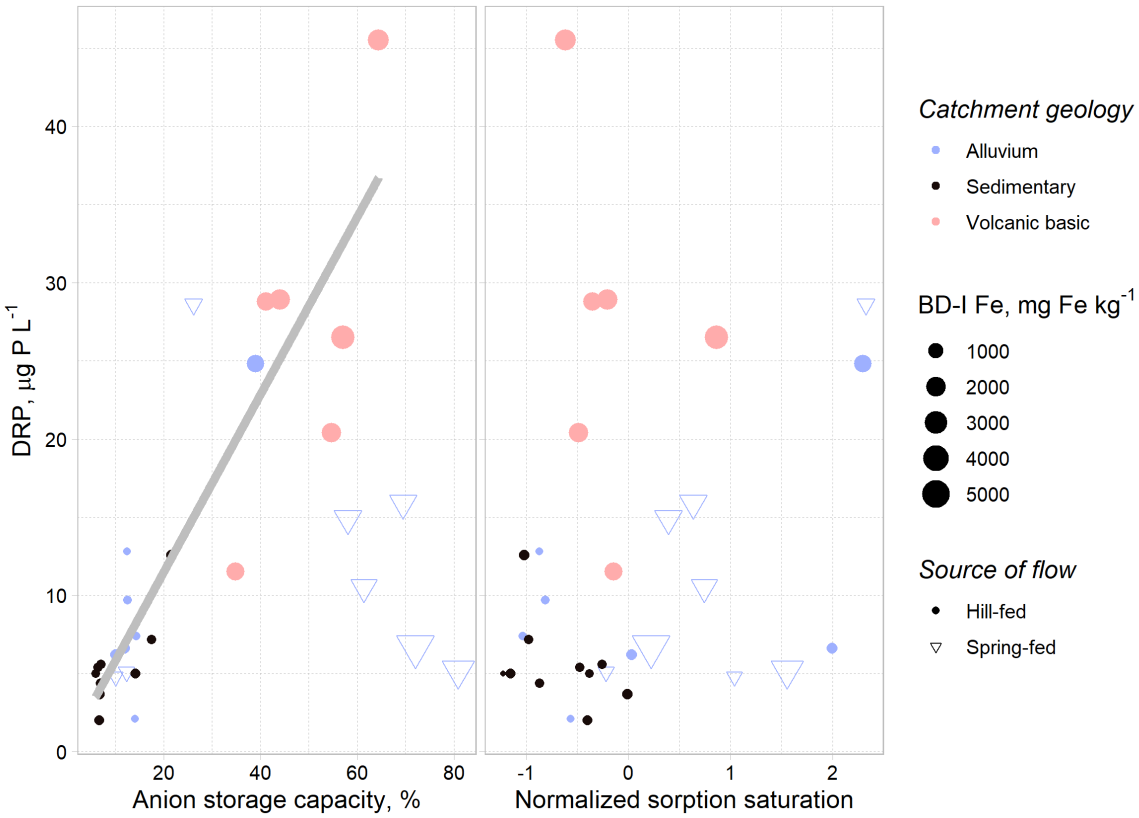

Figure 4.7 In-stream DRP as a function of sediment P sorption metrics: anion storage capacity (0 to 100%) and normalized sorption saturation (scaled value of sediment labile P divided by anion storage capacity; dimensionless). The robust linear model discussed in text is shown.

Contrary to our expectations, sediments with greater sorption potential positively correlated with in-stream DRP (Figure 4.7), but only for the case of hill-fed streams. Using normalized sorption saturation as a proxy for degree of phosphorus saturation of the sediments, there was little relationship between this variable and in-stream DRP. In addition to the linear models fitted with the labile P fractions above, we used ASC as a predictor of DRP – again, we excluded spring-fed sites as a confounding variable. Using ASC alone gave similar performance (RMSE of 4.44 µg P L^{-1}) as the models with terms for geology and the labile P fractions, but with less model degrees of freedom. Overall, the best linear model for DRP employed ASC and H_2O-P (RMSE of 4.38 µg P L^{-1}), with slopes of 2.03 (1.22 to 2.83) µg P L^{-1} per mg H_2O-P kg^{-1} and 0.325 (0.218 to 0.431) µg P L^{-1} per ASC %.

4.5 Discussion

Our data suggests that bioavailable P, i.e., DRP and labile sediment P fractions, in these streams is most likely to be modified by two principal factors: 1) reactive Fe content as derived from both geology and in-stream Fe cycling, and 2) the exchange between sediment porewaters and the water column (i.e., hyporheic exchange) as governed by hydraulic forces and sediment particle size distribution. We expected that the more sorptive sediments, as driven by reactive Fe oxides, would encourage lower DRP concentrations but we found the opposite pattern for streams where hyporheic exchange is more prominent (Figure 4.7). This result has implications for how biogeochemical processes in streams

mediate DRP concentrations and thus influences how we may interpret effects of on-land nutrient pollution management.

4.5.1 Diminished Ca-based mechanisms to buffer DRP in low to moderately alkaline streams

Catchment geology is a primary control on the geochemical composition of stream water (Bluth and Kump 1994) and may promote P attenuation via mineral Ca phase deposition (House 2003). Our results suggest that the Ca-based mechanisms for P removal were minimal in our streams. The stream water column showed no real potential to form hydroxylapatite nor calcite at the time of sampling (see detailed discussion in Appendix D.4), despite the fact that mid-day samples are most likely to capture the peak calcite SI due to photosynthetic activity (Corman et al. 2016). Additionally, the authigenic Ca-P sediment phases (acetate-P) were relatively small (typically <3% total P) and correlated with sediment organic matter concentrations but not with DRP concentrations. The correlation with organic matter suggests that authigenic mineral Ca-P concentrations may be overestimated by the acetate-P fraction (Jensen et al. 1998). However, the acetate-P fractions did slightly improve linear model fits for sediment P sorption after inclusion of amorphous Fe oxides (BD-I Fe; Table D.3), suggesting that associated carbonates in this fraction may have contributed to P sorption capacity. Similarly, Machesky et al. (2010) also determined acetate-P in stream sediments, but for streams more favorable to calcite precipitation (daytime calcite SI of ~1 in some cases): though acetate-P made up little of the sediment P concentration (<5%), authigenic carbonate phases likely increased sediment P sorption alongside Fe oxides. Our snapshot sampling cannot rule out carbonate phases (including calcite) formed on days prior to our sampling, so we are cautious regarding the transient nature of such mineral phases.

Calcium-based mechanisms for P attenuation are important in many streams (Jarvie et al. 2006a; Cohen et al. 2013), but such streams often drain carbonaceous/karst geologies. When the likelihood of Ca mineral interactions – broadly indicated by alkalinity and Ca activity (House 2003; Stets et al. 2017) – is diminished, stream P attenuation may weaken. For example, by experimentally manipulating canopy shading (and thus photosynthetic activity) in calcareous streams (calcite SI of 0.8 to 0.9), Corman et al. (2016) measured a decrease in stream P attenuation due to not only the reduced algal P uptake but also the decreased calcite deposition. For catchments with somewhat less calcareous geology, the potential for calcite deposition is naturally lower. As a result, alternative processes (e.g., sorption) can be relatively more important for buffering DRP concentrations.

4.5.2 Geological and Fe influences on sediment phosphorus

Phosphorus bound in sediments represents a crucial component of legacy P. The sediments themselves are heavily attenuated, where stream networks can retain enormous amounts of sediment and for long residence times (Wohl 2015). Sediment contributions to legacy P, though, are determined by the lability of the various sediment P phases and the potential for sediments to sorb more P. We found that geology greatly determined the predominant forms of sediment P (Figure 4.3; Table 4.4), but most of this

sediment P was unlikely to be bioavailable in our study streams. Taking H_2O and BD-I extractable P (labile P) as potentially bioavailable in lotic systems and definitely available in receiving lentic systems (Golterman 2004; Monbet et al. 2010; Crockford et al. 2015), sediment labile P was generally <10% of total P. Although small relative to more recalcitrant pools, these labile pools are highly reactive under baseflow and may be subjected to microbial turnover (McDowell 2003), exchange with porewaters via desorption or reductive dissolution (Zak et al. 2006; Lewandowski and Nützmann 2010; Loh et al. 2013), and potentially to P-scavenging periphyton mats (Wood et al. 2015).

Catchment geology also influenced sediment Fe concentrations and reactivity. Notably, volcanic basic sediments had relatively greater BD-extractable Fe as did the silty sediments in the spring-fed alluvial streams. The BD-I Fe fraction (poorly crystalline Fe oxides) are particularly reactive towards P due to their high surface areas (Stumm and Sulzberger 1992; Jan et al. 2013, 2015) and this is evident in our data (Table 4.4). Mean BD-I Fe:P ratios (varying with geology from 18.3 to 65.0) were significantly lower than BD-II Fe:P (31.6 to 99.2), owing to the greater affinity for P in the less crystalline fraction. However, the Fe:P ratios for these Fe oxide fractions did not predict sediment P sorption, which contrasts with previous notions that Fe:P indicates available sorption sites (Jensen et al. 1992; Coelho et al. 2004; Kronvang et al. 2009). This could be due to the variable composition of Fe species in these pools (Senn et al. 2015; Herndon et al. 2019). Instead, our data suggests that the mass of BD-I Fe was most predictive for sediment P sorption (Figure 4.6; Table D.3). This result was consistent regardless of sediment texture, source of flow, or catchment geology. Remarkably, despite their high contribution to P sorption capacity, the BD-extractable Fe in these sediments was generally <10% of the total Fe, suggesting that much of the total sediment Fe is noncritical to stream P cycling.

Notwithstanding the contribution towards P sorption of Al oxides (Danen-Louwerse et al. 1993; Mendes et al. 2018) and carbonate minerals, reactive Fe species are strong predictors of P sorption in a variety of aquatic environments (Zhang and Huang 2007; Machesky et al. 2010). For example, Marton and Roberts (2014) and Herndon et al. (2019) found that the Bache-Williams index (BWI) was predicted by amorphous or reactive Fe oxide concentrations in peat, tundra, and marsh soils, much like our findings for benthic stream sediments (Figure 4.6). While these environments differ substantially, a similar theme is apparent regarding Fe-P relationships: where redox interfaces generate amorphous Fe oxides, there is a greater potential for P adsorption. The stability of these sorption sites (e.g., against reductive dissolution; Peiffer et al. 2021) remains a critical topic for P cycling in streams.

4.5.3 Greater sediment P sorption capacity did not decrease DRP concentrations: implications for management

In contrast to what we expected, the greater sediment P sorption capacity as a result of greater reactive Fe oxides did not translate into lower DRP concentrations in streams. Rather, we observed the opposite trend but only for streams with likely sufficient hyporheic exchange (Figure 4.6). We suspect that, given the P sorption potential in these sediments was driven by amorphous Fe oxides (Figure 4.5), the

dynamics of Fe cycling may be coupled with DRP through: 1) dynamic precipitation and dissolution of Fe oxides and their bound P (Runkel et al. 1999; Smolders et al. 2017) and 2) the generation of Fe colloids at redox interfaces which may act as carriers for P and thus elevate DRP (Ren and Packman 2005; Baken et al. 2016b; Gottselig et al. 2017). Both of these potential mechanisms for increasing or maintaining water-column DRP are constrained for streams with little hyporheic exchange (e.g., the spring-fed streams in the present study; Boano et al. 2014). While we have indications of the former (Figure 4.6 and Figure 4.7) and latter (ferrihydrite SI correlation with HPO_4^{2-} activity; Figure D.2) hypotheses, more investigation would be needed to isolate the responsible processes.

Recently, Dupas et al. (2017) observed elevated DRP concentrations in forested streams during summer low flow conditions despite no presence of point sources. They hypothesized that Fe colloids generated in the waterlogged riparian soils could carry P and promote greater DRP concentrations, similar to the relationship we found for our study streams. These results point towards the need for understanding the role of Fe cycling in controlling P cycling in streams and therefore in identifying possible sources of P pollution.

It is confounding that greater sediment P sorption capacity did not necessitate lower DRP concentrations. For streams at baseflow, we generally expect that greater P sorption leads to both lower DRP concentrations (Haggard et al. 2007; McDaniel et al. 2009; Weigelhofer et al. 2018a) and fluxes (Baulch et al. 2013). This is typically accurate for soils, where more sorptive soils (e.g., greater ASC) leads to greater retention of P (McDowell et al. 2015). The comparison to soils is an important one, since what we may understand about P cycling in soil environments may not generalize to stream sediments. Critically, while nutrient pollution management can target soils to mitigate legacy P, the same cannot be said for sediments (McDowell et al. 2004; Macintosh et al. 2018; Kusmer et al. 2019).

Further research is needed to link the abiotic P exchange mechanisms, alongside biotic P cycling, to the spatiotemporal DRP signal observed in the water column. Reach-scale studies of P attenuation (Ensign and Doyle 2006) are valuable for documenting the extent of P attenuation but are unspecific regarding which processes in streams ultimately buffer P pollution. While stream sediments are well-known to provide reaction sites for P, little has been done to link these reaction sites in streams to P cycling in a mechanistic manner, i.e., by connecting these zones of reactivity to hydrological transport (Boano et al. 2014). Broadly, abiotic P cycling is tied to Ca and Fe biogeochemical cycles whose relative importance depends on stream and catchment properties (e.g., geology). We suggest that future research on stream P cycling incorporates the multiple mechanisms – both biotic and abiotic – where relevant. With detailed understanding, we may be able to address the problem of legacy P and how to better mitigate P pollution of our waters.

Chapter 5

The biotic contribution to the benthic stream sediment phosphorus buffer

5.1 Abstract

Benthic stream sediments interact strongly with phosphorus (P) and can buffer dissolved reactive P (DRP) concentrations. The sediment P buffer can be measured with the sediment equilibrium phosphate concentration at net zero sorption (EPC_0), which often correlates well with DRP. Yet, it is unclear how much of this P affinity in sediments is attributable to biotic (microbial P demand) or abiotic (sorption) processes. To clarify the role of biotic processes on EPC_0, we used two experiments with benthic sediment from 12 streams. First, sediments sterilized by γ-irradiation increased in EPC_0 compared to fresh sediments by a median of 83%. This increase in EPC_0 was likely a result of cell lysis, where microbial biomass P (2.4 to 22.6 mg P kg^{-1}) was re-adsorbed to sediment surfaces. This data also shows that the sediment microbial biomass is a significant, yet under-reported biotic stock of P in streams compared to their photic zone counterpart (i.e., periphyton). In a second experiment, fresh sediment EPC_0 was measured after alleviating potential limitation of carbon (C) and nitrogen (N) for microbial growth. Sediment EPC_0 did not change with C addition and decreased slightly (0.5 µg P L^{-1} or ~5% decrease) with N addition, suggesting these sediments strongly buffered DRP towards the EPC_0 in spite of biotic demand. Together, these experiments suggest that sediment EPC_0 was primarily abiotic in nature but that sediments may subsidize biotic P requirements through desorption. Further work is needed on whether this relation holds for streams with different substrate, geology, and nutrient inputs.

5.2 Introduction

Benthic sediments provide much of the phosphorus (P) attenuation observed in streams (Haggard and Sharpley 2007; Hamilton 2012). Sediments contain inorganic particles of varying sizes and geochemical characteristics derived from the parent material, which determine the sites available for P sorption (House 2003). Sediments also contain allochthonous and autochthonous stocks of organic matter (OM; Tank et al. 2010; Kaplan and Cory 2016), providing fuel for the metabolism of sediment microbial biofilms (Battin et al. 2016). These characteristics make the sediment matrix a hotspot of biogeochemical cycling for P as well as for carbon (C) and nitrogen (N). Hence, when P is removed from the water column at baseflow, it is difficult to pinpoint which biotic and abiotic mechanisms are responsible. Such mechanisms, when known, could inform better nutrient pollution modelling (Macintosh et al. 2018) and subsequent strategies to mitigate the effects of P enrichment (Meals et al. 2010; McDowell et al. 2018; Drohan et al. 2019).

The range of abiotic and biotic mechanisms that remove dissolved reactive P (DRP) from solution and into sediments is diverse. For example, reactive surfaces on particles, such as highly sorptive hydrous metal oxides (Dzombak and Morel 1987; Sposito 2004), often constitute much of the abiotic P sorption capacity and P storage in soils and sediments (Small et al. 2016; Audette et al. 2018; Herndon et al. 2019). Meanwhile, microorganisms entrained in the sediment matrix also require P for their growth and likely supply much of the biotic P uptake for stream sediments (Sinsabaugh et al. 2009; Horn et al. 2011; Hill et al. 2012). This biotic P uptake is dependent on stoichiometric demand (Cross et al. 2005; Maranger et al. 2018). While microorganisms beneath the benthic zone (primarily heterotrophs due to lack of light; Battin et al. 2016) can vary in stoichiometric P requirements (Cross et al. 2005; Scott et al. 2012), they approximate a molar C:N:P ratio of 60:7:1 (Cleveland and Liptzin 2007; Hill et al. 2010; Sinsabaugh et al. 2012). This ratio is P-rich relative to that of other lotic biota (Cross et al. 2005) and indicates the potential for sediment microorganisms to moderate stream P cycling if sufficient C and N are available. The distinction between abiotic and biotic sediment P removal is important, however, as not only are their rates and capacities different, they also differ in their responses to other processes within the stream (e.g., varying C and N resource supplies or changes in pH).

Sediments can strongly buffer solution DRP towards a concentration termed the equilibrium phosphate concentration at net zero sorption (EPC_0; Froelich 1988). The EPC_0 is measured from a series of batch sediment incubations with solutions of varying initial phosphate concentrations; the concentration where neither net removal nor release of P by the sediment occurs (i.e., the *x*-intercept from a P sorption plot) is the EPC_0 (Taylor and Kunishi 1971; Froelich 1988). As such, sediment EPC_0 relates to the potential of a sediment to buffer variable DRP concentrations in the water column. For example, Ekka et al. (2006) showed increases of several mg P L^{-1} in sediment EPC_0 downstream from point source inputs of DRP relative to sediments immediately upstream, showing how the EPC_0 is elevated with greater loadings of P. In large stream survey studies, McDaniel et al. (2009) and McDowell (2015) found that

EPC_0 not only related to P loading (e.g., relatively greater when influenced by point sources or agricultural land use) but also correlated with particle size and geochemical characteristics. That is, given a similar P loading, more (chemically) sorptive sediments will have a lower EPC_0 and thus a greater potential to buffer in-stream DRP concentrations.

EPC_0 does not differentiate between abiotic or biotic mechanisms that control sediment P flux. The sediments used for EPC_0 likely harbor both abiotic and biotic P exchange processes. So, a recurring question for stream P cycling has been: What is the relative contribution of abiotic and biotic processes towards the EPC_0? By sterilizing the sediments via varying methods, EPC_0 has been observed to increase dramatically or by very little, with similarly inconsistent changes in P sorption capacity (Meyer 1979; Klotz 1985; Munn and Meyer 1990; Haggard et al. 1999; Lottig and Stanley 2007; McDaniel et al. 2009; Griffiths and Johnson 2018). Unfortunately, such comparisons can suffer from the choice of sterilization method, as several methods cause severe physicochemical changes to the sediment. For example, autoclaving can disrupt soil OM and surface chemistry (Trevors 1996; Buessecker et al. 2019). Measurements of abiotic P sorption will not be representative unless sediment physicochemistry is undisturbed. Further, few EPC_0 comparisons have included corrections for cell lysis following sterilization, where microbial biomass P can release into solution and may adsorb onto the sediment surface (e.g., Klotz 1985; McDaniel et al. 2009).

This study seeks to distinguish biotic and abiotic contributions to stream benthic sediment EPC_0 while minimally disrupting sediment physicochemistry. We account for P in the microbial biomass and determine if removing C- and N-limitation on biotic P demand affects the EPC_0. Our two hypotheses were that (1) abiotic processes determine EPC_0 and so EPC_0 would be similar for fresh and sterilized sediments if microbial biomass P were negligible and that (2) microbial P demand, and so the biotic contribution to sediment EPC_0, would be most pronounced under P-limitation vs. C- or N-limitation. For the first experiment, we measured EPC_0 in sediments both fresh and following the use of γ-irradiation to sterilize sediments. For the second experiment, we measured sediment EPC_0 with and without addition of C and N to remove microbial C- and/or N-limitation.

5.3 Methods

5.3.1 Sites and sampling

We sampled 12 streams across Banks Peninsula, New Zealand (Figure 5.1), thus targeting streams draining a consistent geology (volcanic-basic; Snelder and Biggs 2002) but a variety of land uses, and therefore, variable C, N, and P inputs (Table E1). The streams were 2nd to 4th order, with catchment areas ranging from 3.7 to 51 km^2. Benthic substrates were generally gravel to gravel/cobble. Catchment land

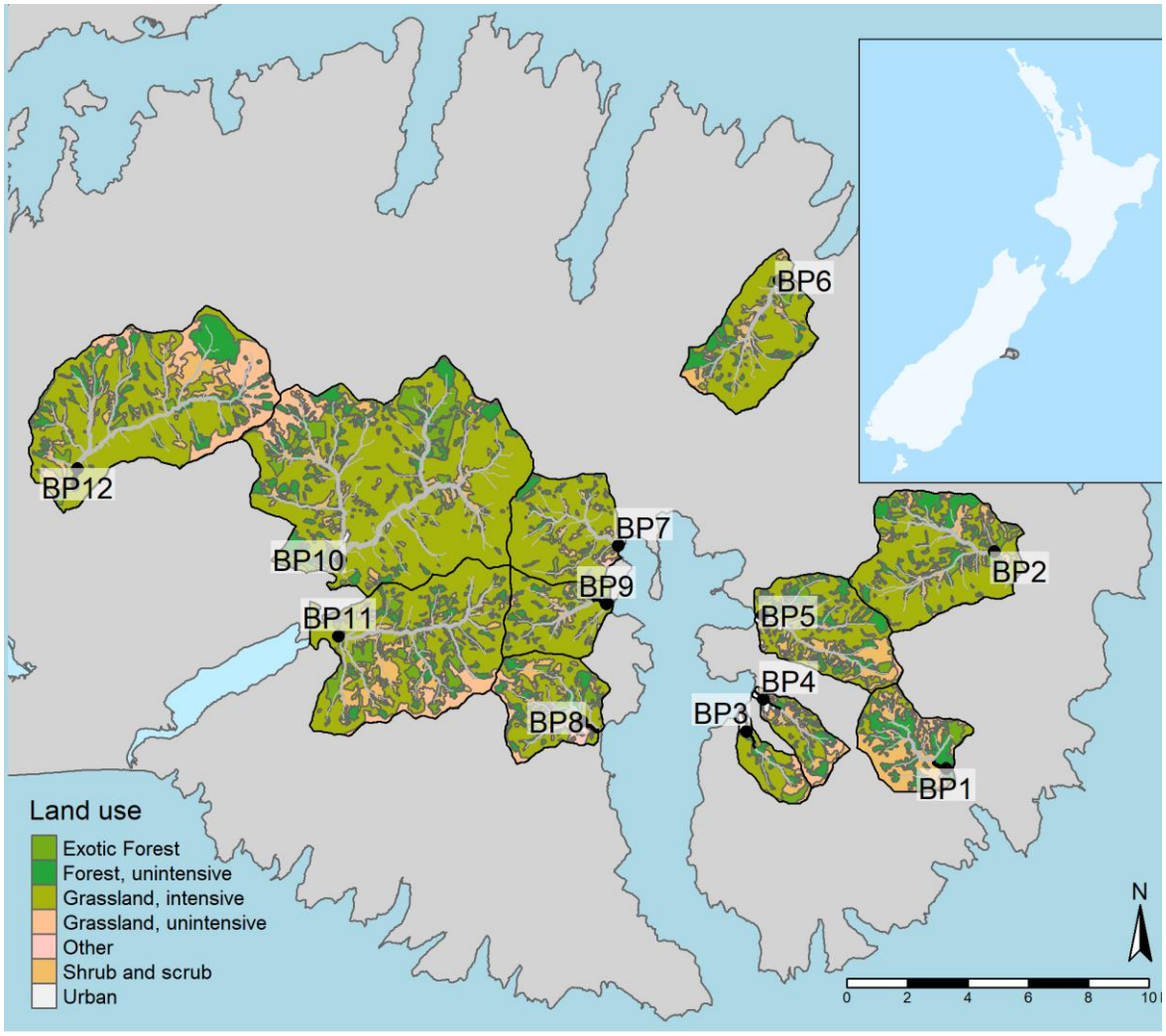

Figure 5.1 Study streams on Banks Peninsula, New Zealand. Detailed site information is given in Table E1. Land uses for 2018 were simplified from the LCDB v5.0 database

uses for 2018, derived from the New Zealand Land Cover database (v5.0; https://lris.scinfo.org.nz/), varied from predominantly native forest and shrub (83%) to predominantly intensive grassland (82%).

Our sampling was designed to accommodate an immediate delivery of sediments to be γ-irradiated with minimal storage time. Hence, we sampled 7 sites one day prior to shipping and five sites on the day of shipping samples for γ-irradiation.

All streams were sampled at summer baseflow conditions (November 2019) with no recent disturbances. Benthic sediments were collected in actively flowing zones, primarily in riffles (i.e., avoiding depositional pools and lateral storage areas). To collect enough mass and to capture the spatial variability in sediments across the channel, we sampled ~10 to 30 m of stream length by sampling in a zig-zag pattern. We collected the top 1-5 cm of substrate with a shovel, wet-sieved the material (<2 mm), decanted excess water after settling, and kept the composite, fine sediment sample (approximately 0.5 to 1 kg per site). A water sample was collected from the thalweg and filtered (0.45 µm) into vials with minimal headspace. Both sediment and water samples were kept on ice and in the dark until they were refrigerated (4 °C) in the laboratory and a set of subsamples for stream water were frozen (-20 °C).

Additionally, we used a HACH HQ40D meter to measure stream temperature, dissolved oxygen (via an optical probe), pH, and specific conductivity *in situ*; all probes were calibrated on the day of sampling according to manufacturer instructions.

5.3.2 Sterilization via γ-irradiation

To control for potential microbial P uptake, we sterilized sediment subsamples via γ-irradiation, as this method non-invasively kills or damages cells with minimal effects on sediment physicochemistry (McLaren 1969; Trevors 1996; Buessecker et al. 2019). Subsamples from all 12 sediments were γ-irradiated with a nominal 25 kGy dose from a ^{60}Co source (MSD Animal Health, Upper Hutt, New Zealand). The actual dose delivered was later confirmed to be 26 kGy (Dave Harris, MSD Animal Health, pers. comm.). This dose is comparable to that used in previous studies and represents a favorable tradeoff between potential physicochemical disruption of the sediment and effective sterilization (Meyer 1979; Östlund et al. 1989; Qiu and McComb 1995; McNamara et al. 2003; Buessecker et al. 2019).

We managed the logistics of sampling, sterilization, and analyses so that (1) sediments were most reflective of *in situ* conditions for EPC_0, (2) sterilized sediments had minimal time for microbes to re-proliferate, and (3) both fresh and sterilized sediments were held under similar conditions throughout prior to analyses. Sediment samples were immediately shipped to be γ-irradiated following collection. Samples were kept cool throughout all handling via freeze-packs (with replacement as needed), thus maintaining temperature comparable to that of the fresh sediments (kept refrigerated in the laboratory at 4 °C). Once γ-irradiated, sterilized sediments were returned to the laboratory within 24 h (4 days since sampling). Using aseptic handling for the sterilized sediments, analyses for dehydrogenase activity (DHA; both γ-irradiated and fresh sediments) and for EPC_0 (γ-irradiated) began immediately (see below). All batch incubations for γ-irradiated sediment EPC_0 were completed within 48 h of sterilization.

5.3.3 Water analyses

Dissolved reactive P was measured on filtered stream samples within 24 hours. We used the molybdenum-blue method (Murphy and Riley 1962) with a 5 cm quartz cell (method detection limit of ~2 μg P L^{-1}). Replicate measurements and external quality-control P standards established relative error at <5%. Frozen water samples were thawed and immediately analyzed for dissolved organic C (DOC) via a TOC analyzer and for mineral N via flow-injection analysis. Two samples were below detection for NO_3-N (0.1 mg N L^{-1}) while all but one sample (0.13 mg N L^{-1}) were below detection for NH_4-N (0.1 mg N L^{-1}).

5.3.4 Sediment physicochemical analyses

Throughout, we refer to wet, un-sterilized sediments as 'fresh' and wet, sterilized sediments as γ-irradiated. Except where noted, analyses on these sediments are for wet sediments but all values are given on a dry-weight (d.w.) mass basis.

A subsample of sediment was oven-dried (105 °C) overnight to measure moisture content. Another subsample was freeze-dried for total elemental and particle size analyses (below) as well as for storage. For the following sediment analyses, we used either duplicate or triplicate measurements. Sediment pH was measured on fresh sediments in DI H_2O with a 1:5 g g^{-1} (sediment:solution) after 30 minutes equilibration. Water extractable P was measured by shaking fresh sediments with DI H_2O (1:10 g g^{-1}) for one hour, centrifuging (2400 *g* for 10 minutes), filtering the supernatant (0.45 µm) and analyzing for DRP as above. Anion storage capacity (ASC; Saunders 1965) of sediments was measured by shaking 1 g d.w. of fresh sediment in 5 mL of a 1000 mg P L^{-1} solution (as KH_2PO_4) in an acetate buffer (pH adjusted to 4.60) for 24 h. The remaining DRP in this extract was measured and ASC was expressed as % of the original concentration removed. Sediment particle sizes (on percent volume basis; Eshel et al. 2004) were determined with laser-diffraction (<1 mm fraction only) with a Malvern Mastersizer 2000 particle size analyzer according to Sperazza et al. (2004). To meet instrumental constraints for particle size analyses, freeze-dried sediments (un-sterilized only) were sieved to <1 mm, which represents the majority of the fine sediments collected here. Sediment total C and N were measured on freeze-dried sediments via dry-combustion (Carter and Gregorich 2007) with an Elementar Vario-Max CN elemental analyzer. Sediment total P and metals (Al, Ca, Fe, and others) were measured on freeze-dried sediments via ICP-OES (Varian 720-ES) following a microwave digestion with nitric acid and hydrogen peroxide (Campisano et al. 2017; US EPA 2019 method 3050B).

5.3.5 Sediment microbial enzymes and biomass P

We measured dehydrogenase activities (DHA) on both fresh and γ-irradiated sediments as an indicator for 1) microbial metabolic activity in fresh sediments and 2) sterile conditions for γ-irradiated sediments prior to EPC_0 incubations. DHA relates to the breakdown of organic compounds during microbial respiration in soils and sediments (Hill et al. 2002, 2012; Prosser et al. 2011). DHA should be minimal for the γ-irradiated sediments since γ-irradiation eliminates or inactivates most microbes and enzymes (McLaren 1969; Tabatabai 1994). However, we note that some enzymes may persist after γ-irradiation (Powlson and Jenkinson 1976; Blankinship et al. 2014) and the DHA method used here (below) may give false positive readings for the sterilized samples since γ-irradiation can reduce minor amounts of redox-sensitive species (Östlund et al. 1989; Buessecker et al. 2019). Similarly, preliminary tests with autoclaved sediments yielded only a ~75% decrease in DHA relative to fresh sediment DHA. While not perfect, we expected DHA to be greatly reduced for sterilized sediments as a means to ensure that negligible sediment microbial activity occurred during EPC_0 experiments.

For the DHA assay, we used 2,3,5-triphenyltetrazolium chloride (TTC) as the substrate, which is reduced enzymatically to triphenylformazan (TPF) as the product (Tabatabai 1994; Öhlinger and Von Mersi 1996). Two g d.w. of fresh and γ-irradiated sediments were incubated with two mL of a 0.5% (w/v) TTC solution – as recommended for coarse sediments with relatively low organic matter – in a 0.1 M TRIS buffer (pH adjusted to 7.6) for 24 h at 25 °C. The incubation was terminated by adding 10 mL of methanol (AR grade; ≥99.8%) and vortexing. After centrifuging, we analyzed the supernatants

by measuring absorbances at 485 nm within 1 h with a 1 cm light path. Standards of TPF were prepared in methanol and analyzed in the same fashion. Method blanks were included throughout and sediments were analyzed in triplicate or duplicate (if sample amount was limited). We took care to minimize exposure to light during handling (Öhlinger and Von Mersi 1996). DHA is expressed here as mg TPF kg^{-1} h^{-1}.

We estimated microbial biomass P in the fresh sediments following the methodology of Brookes et al. (1982) and McLaughlin et al. (1986) as outlined for sediments by McDowell (2003). Notably, we analyzed sediments wet rather than dry, as recommended for soil microbial biomass P analysis by Brookes et al. (1982). Additionally, while clear guidance is lacking on storage times, microbial biomass measurements for soils kept cool are stable for several weeks following initial disturbance effects from sampling (Kouno et al. 1995; Turner and Romero 2010). Hence, our analyses began ~3 weeks after sampling. Briefly, 1 g d.w. of fresh sediment was weighed into three centrifuge tubes, in duplicate. To one (P_{killed}), 0.5 mL of liquid chloroform stabilized in amylene was added (rather than as vapor; McLaughlin et al. 1986), capped, and mixed while the other two treatments (P_{fresh} and P_{spike}) were capped – all tubes then incubated at room temperature for 24 h. Each treatment was then extracted with 0.5 M $NaHCO_3$ (pH adjusted to 8.5) for 30 minutes, with the P_{spike} treatment receiving an additional P spike in the extraction equivalent to 25 mg P kg^{-1} sediment. The reactive P in the extracts was measured with a modified molybdenum-blue method suitable for alkaline extracts (Dick and Tabatabai 1977; He and Honeycutt 2005). We note that the P_{fresh} treatment is equivalent to Olsen P, an indicator for bioavailable P. Microbial biomass P (mg P kg^{-1}) was calculated by:

$$\frac{(P_{killed} - P_{fresh})}{K_P} \times \frac{spike}{(P_{spike} - P_{fresh})}$$

where all P values (P_{killed}, P_{fresh}, P_{spike}, and *spike*) are given on a mg P kg^{-1} basis and K_P is a coefficient for the recovery of microbial biomass P with chloroform killing. Here, we assumed K_P to be 40% (Brookes et al. 1982; McLaughlin et al. 1986; Jenkinson et al. 2004).

5.3.6 Equilibrium phosphate concentrations at net zero sorption (EPC_0) and nutrient treatments

The equilibrium phosphate concentration at net zero sorption (EPC_0) was determined for γ-irradiated sediments and then fresh sediments in close succession (completed within 48 and 96 h of sterilization, respectively). We weighed 0.5 g d.w. of wet sediment (Simpson et al. 2019) into 15 mL centrifuge tubes and added 10 mL of solution (DI water with background of 3 mM $CaCl_2$; Lucci et al. 2010) adjusted to four P concentrations (as KH_2PO_4): 0, 30, 100, and 250 µg P L^{-1}. Up to three replicates were used at each concentration depending on the amount of sample available. The tubes were shaken (end over end) for 16 h, centrifuged, and supernatants filtered (0.45 µm). Extracts were then refrigerated until analyzed for DRP within 24 h. Standards for DRP were prepared in the same background matrix (3 mM $CaCl_2$).

For the nutrient addition experiment, we alleviated C- and/or N-limitation during the fresh sediment EPC_0 incubations by adding labile C and/or N to the solutions in a factorial experimental design. We targeted a molar C:N:P ratio of 200:20:1 to create an environment enriched in C and N but limiting in P. Nutrient treatments were +C (C:P of 200:1), +N (N:P of 20:1), C+N (C:N:P of 200:20:1), or 'none' (only P). We added C as D-(+)-glucose and N as KNO_3 to the solutions used for EPC_0. We also applied the C+N treatment to the γ-irradiated sediment as a check but these data yielded no discernable difference from the standard EPC_0 treatment and so are not discussed for sake of brevity.

Each sediment EPC_0 was then determined as the *x*-intercept from regressing P sorption (mg P kg^{-1}) against initial P concentration (Simpson et al. 2019; Appendix A). We excluded the points from the 250 µg P L^{-1} treatment, as 100 µg P L^{-1} already provided suitable adsorption points and greater P concentrations only diminishes the linearity of the observed P sorption near the EPC_0. We estimated the uncertainty in the measured EPC_0 with a 95% confidence interval.

Further, under our alternative hypothesis that microbial biomass P can shift EPC_0 following sterilization, we predicted changes in EPC_0 post-sterilization assuming that all microbial biomass P would be re-adsorbed upon lysis. To this end, we calculated the linear sorption slopes on an equilibrium P concentration basis and projected these slopes according to the microbial biomass P; i.e., we divided microbial biomass P by the equilibrium sorption slope to get the expected change in EPC_0. This result would be the expected increase in EPC_0 if 100% of microbial biomass P were adsorbed by the sediments while assuming nearly linear adsorption rates. We compared these predictions to the observed changes in EPC_0 following γ-irradiation.

5.3.7 Statistical analyses

Statistical summaries reported here are generally the median (mean ± SD). However, we do not summarize NH_4-N data (too many missing values) and summarize NO_3-N data (two out of 12 values below detection limit) via methods for censored data (Helsel 2005). For NO_3-N summaries, we apply the Kaplan-Meier method as implemented in the 'NADA' R package (Lee 2020).

For the sterilization experiment (a paired treatment design), we analyzed the changes in EPC_0 (n=12) with a paired Wilcoxon signed rank test. The null hypothesis is that the distribution of pairs has a median of 0 (Hollander et al. 2013a). For the alternative case, we estimated the shift in the median with the pseudo-median and its 95% confidence interval. We also examined changes in EPC_0 following γ-irradiation with Spearman correlations versus catchment, sediment, and stream variables.

For the nutrient addition experiment, we sought to test for the effect of nutrients (+C and +N) and their interaction (C+N) on EPC_0 via a two-way layout (n=48). The facts that 1) EPC_0 were not estimated with the same certainty (i.e., varying confidence intervals) and 2) that the EPC_0 were dependent upon the sediment analyzed (i.e., there are 12 'clusters' in the data) meant that this information needed to be incorporated in our test. Thus, we analyzed this experiment with a mixed-effects model, where sediment

could be a grouping variable (random effect). We used nonlinear mixed-effects (R package 'nlme'; Pinheiro and Bates 2000; Pinheiro et al. 2020) so that we could re-write the linear sorption model to have the *x*-intercept (the EPC_0) as a parameter to be estimated and to test for nutrient effects on EPC_0 directly, making greater use of the data. More details on this approach are in Appendix E.2.

All analyses were conducted in R, ver. 4.0.2 (R Core Team 2020).

5.4 Results

5.4.1 Stream and sediment characteristics

At time of sampling, the streams reflected typical summer baseflow conditions (Table 5.1). While the streams had appreciable DOC (median, 1.73 mg L^{-1}) and DRP concentrations (49 µg L^{-1}), NH_4-N concentrations were mostly below detection (one stream, BP9, measured 0.13 mg L^{-1}) and NO_3-N concentrations were also low (0.21 mg L^{-1}). Using DOC and NO_3-N, stream water molar C:N had a median of 10.4. The fine (<2 mm) benthic sediments (Table 5.2) were sandy, but quite sorptive (median ASC of 56% in a range of 0-100%), as expected for sediments with volcanic geology. Sediment organic matter content varied considerably across the sediment from 12 streams, with mean (± SD) sediment total C and total N concentrations of 12.8 (± 7.1) and 0.86 (± 0.51) g kg^{-1}, respectively. Median sediment C:N was 16.9. Correlation matrices and Spearman correlations for select variables are shown in supplementary Figures E1-E3 (see Appendix E). No strong correlations between catchment land use and stream or sediment physicochemical characteristics were noted.

Table 5.1 Stream water (*n*=12) physicochemistry (at time of sampling), dissolved organic C (DOC), nitrate-N (NO_3-N), and dissolved reactive P (DRP). Summary values are given as the median (mean ± standard deviation). DO is dissolved oxygen.

Variable	Unit	Summary
Temperature	°C	13.7 (13.8 ± 1.5)
DO	% saturation	99.4 (99.6 ± 3.3)
Specific conductivity	µS cm^{-1}	148 (147 ± 16)
pH	S.U.	7.58 (7.54 ± 0.18)
DOC	mg L^{-1}	1.73 (1.73 ± 0.67)
NO_3-N	mg L^{-1}	0.195 (0.200 ± 0.078)
DRP	µg L^{-1}	48.6 (48.9 ± 8.7)

Table 5.2 Benthic sediment (n=12 streams) characteristics, total elemental concentrations, extractable P, and sorption as anion storage capacity. Summary values are given as the median (mean ± standard deviation).

Variable	Unit	Summary
pH	S.U.	7.03 (6.99 ± 0.149)
sand	%	79.9 (77.3 ± 9.4)
silt	%	17.8 (20.2 ± 8.11)
clay	%	1.95 (2.49 ± 1.32)
Total C	g kg^{-1}	11.2 (12.8 ± 7.05)
Total N	g kg^{-1}	0.833 (0.861 ± 0.513)
Total P	g kg^{-1}	2.18 (2.17 ± 0.48)
Total Fe	g kg^{-1}	58.4 (61.9 ± 8.6)
Total Al	g kg^{-1}	41.8 (43.0 ± 9.6)
Total Ca	g kg^{-1}	8.90 (9.01 ± 1.57)
Water extractable P	mg kg^{-1}	4.15 (4.38 ± 1.22)
Olsen P	mg kg^{-1}	26 (27.6 ± 5.1)
Anion storage capacity	%	55.6 (53.5 ± 8.06)

5.4.2 Gamma irradiation experiment

Fresh sediment dehydrogenase activities (DHA) varied widely (Figure 5.2), with mean activities of 1.89 (± 0.82) mg TPF kg^{-1} hr^{-1}. Fresh sediment DHA was closely associated with sediment total C (ρ =0.91, p <0.001) and total N (ρ =0.93, p <0.001). However, DHA decreased for all samples following γ-irradiation (62 to 86% decrease), commensurate with previous experimental work for soils (McNamara et al. 2003; Gebremikael et al. 2015). We consider the γ-irradiated samples here to have had negligible microbial activity throughout the EPC_0 measurements given that (1) the decreasing trend in DHA with

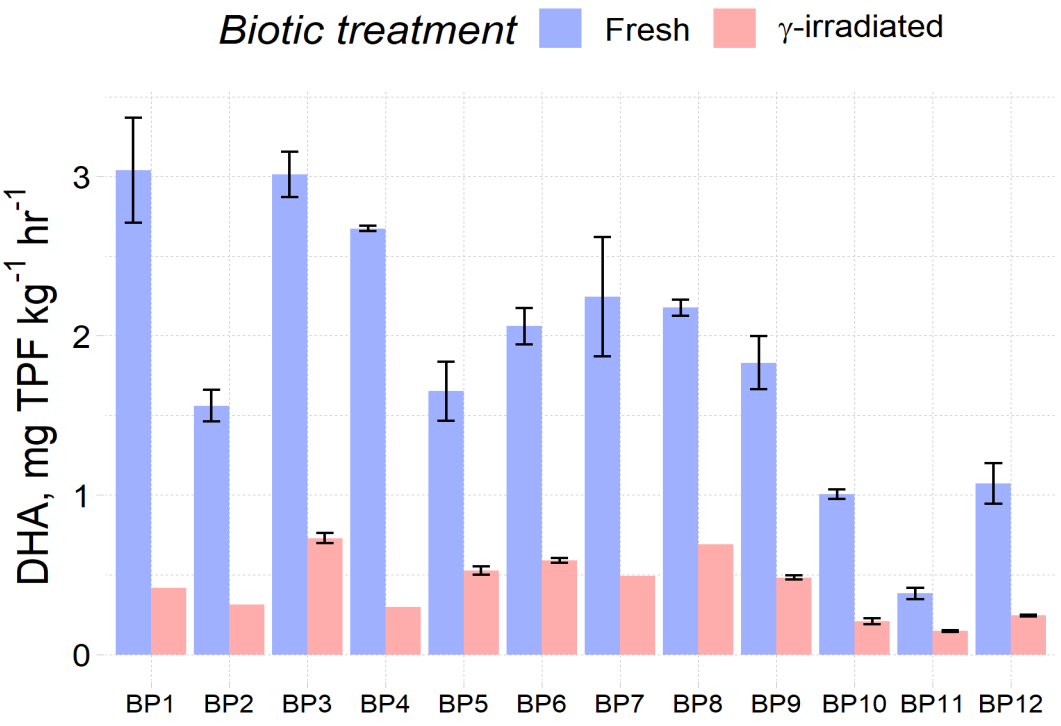

Figure 5.2 Dehydrogenase activities (DHA) in sediments from all 12 streams and both microbial treatments measured at the same time (post γ-irradiation). For fresh sediments, three replicates were measured and the means are shown with standard errors. However, due to limited sample amount for some γ-irradiated sediments, standard errors could not be calculated for sterilized sediments 1, 2, 4, 7, and 8.

sterilization was consistent and (2) the irradiation dose given (26 kGy) was sufficient to eliminate the majority of microbes and inhibit growth for the timeframe of our experiment (48 h post sterilization; McLaren 1969; Östlund et al. 1989; McNamara et al. 2003).

Fresh sediment EPC_0 ranged from 7.6 to 14.8 µg P L^{-1} (Figure 5.3) and correlated with in-stream DRP concentrations (ρ =0.65, p =0.026). Sediment microbial biomass P averaged 11.5 (± 5.6) mg P kg^{-1} and correlated closely with fresh EPC_0 (ρ =0.87, p =3.1e-4).

Following γ-irradiation, EPC_0 increased for all sediments studied, though 95% confidence intervals for the calculated EPC_0 overlapped for one sediment (BP2; Figure 5.3). A paired Wilcoxon signed rank test estimated a median increase in EPC_0 of 8.32 (6.38 to 12.6) µg P L^{-1} following γ-irradiation, an 83% increase relative to the median fresh EPC_0. Notably, this consistent increase in EPC_0 with sterilization relative to fresh EPC_0 was related to microbial biomass P (Figure 5.4a; ρ =0.63, p =0.032), with a regression slope of 0.635 (0.268 to 1.00) µg P L^{-1} per mg P kg^{-1} (the intercept was not different from zero). We note no other clear effect of γ-irradiation on P sorption relative to that of fresh sediments (e.g., changes in sorption slopes or variability in sorption).

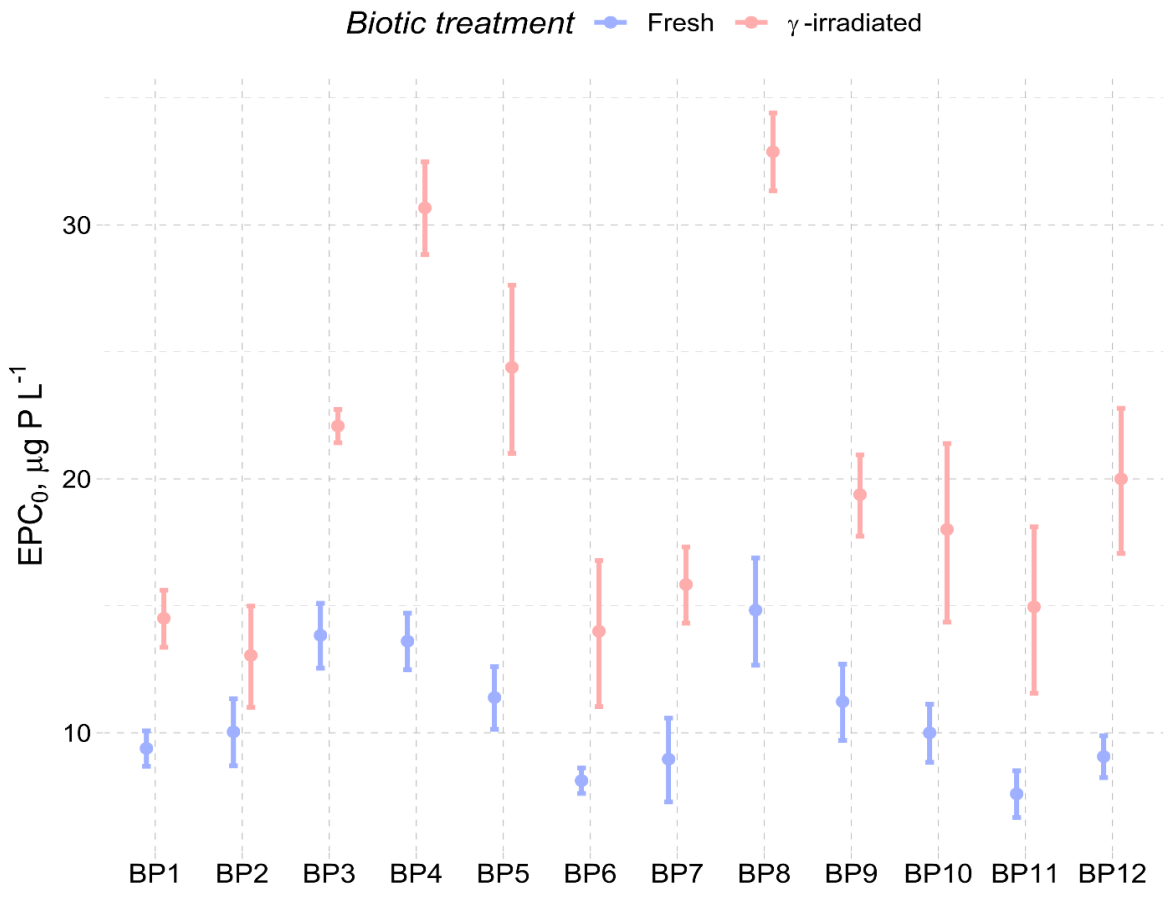

Figure 5.3 Effect of γ-irradiation on sediment equilibrium phosphate concentrations at net zero sorption (EPC_0). Bars indicate the 95% confidence interval.

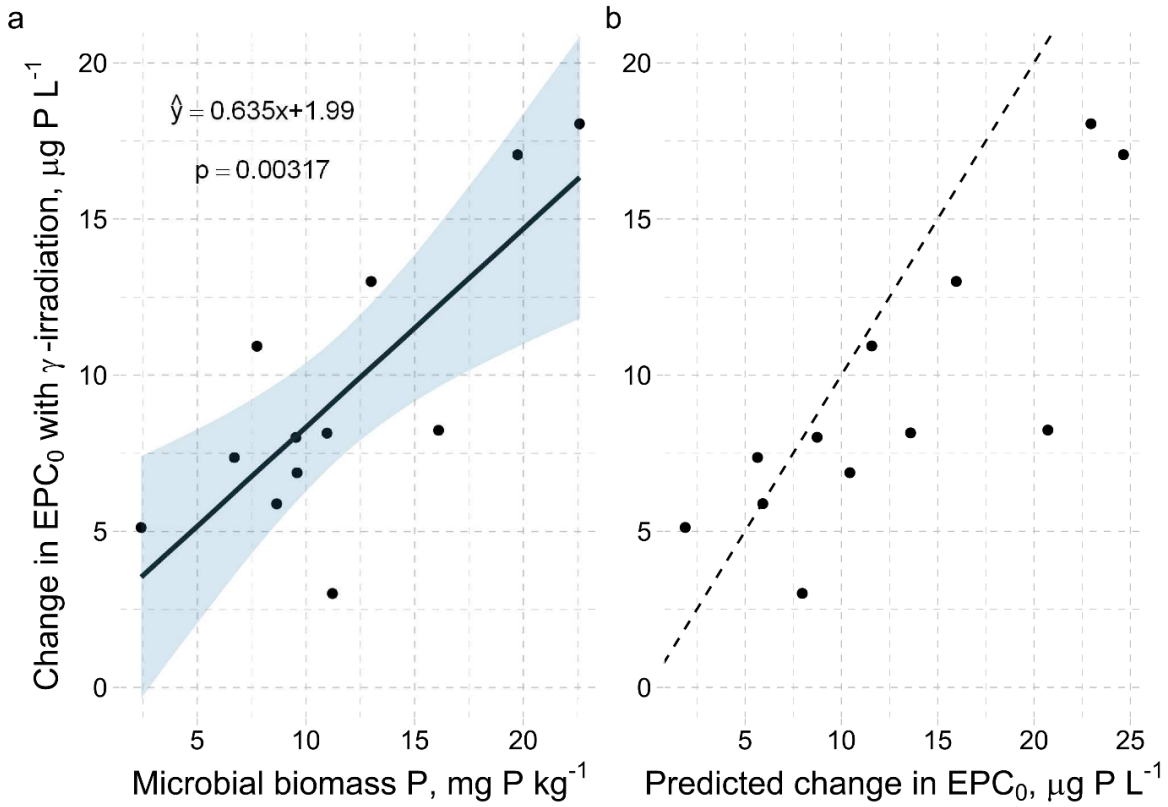

Figure 5.4 The increase in sediment EPC_0 due to γ-irradiation (sterilized EPC_0 minus fresh EPC_0) plotted (a) as a function of sediment microbial biomass P. A linear regression with standard error about the fit is shown. This change in EPC_0 with sterilization is also plotted (b) against the predicted change in EPC_0 based on the adsorption data and the microbial biomass P assuming 100% adsorption of the lysed P (dashed line is the 1:1 line). Note that these estimates could be negatively biased since we extrapolated a linear sorption curve to equilibrium concentrations past the likely range of linear sorption.

We examined the relationship between changes in EPC_0 post-sterilization and microbial biomass P under the hypothesis that shifts in EPC_0 may be due to adsorption of the P flush following cell lysis. The predicted increases in EPC_0 based on the adsorption of the microbial biomass P are plotted against actual changes in EPC_0 following γ-irradiation in Figure 5.4b. These independent estimates were consistent with the measured changes in EPC_0 post-sterilization (root mean square error (RMSE) of 5.22 µg P L^{-1}) but were positively biased for larger magnitude changes.

5.4.3 Nutrient amendment experiment

Fresh sediments were analyzed for EPC_0 with additions of glucose and/or KNO_3 solutions to remove C- and/or N-limitation on microbial growth. When considering the measurement uncertainties within a given sediment, EPC_0 did not appear to vary significantly between the nutrient treatments for any one sediment (Figure 5.5). However, when modeling the full data set with nonlinear mixed effects, we found a statistically significant decrease in EPC_0 with N addition but no significant change with C addition (Table 5.3; more modeling details in Appendix E). The likelihood-ratio tests and Akaike Information Criterion showed that more complex models with an effect for +C (p=0.07) and the interaction effect (C

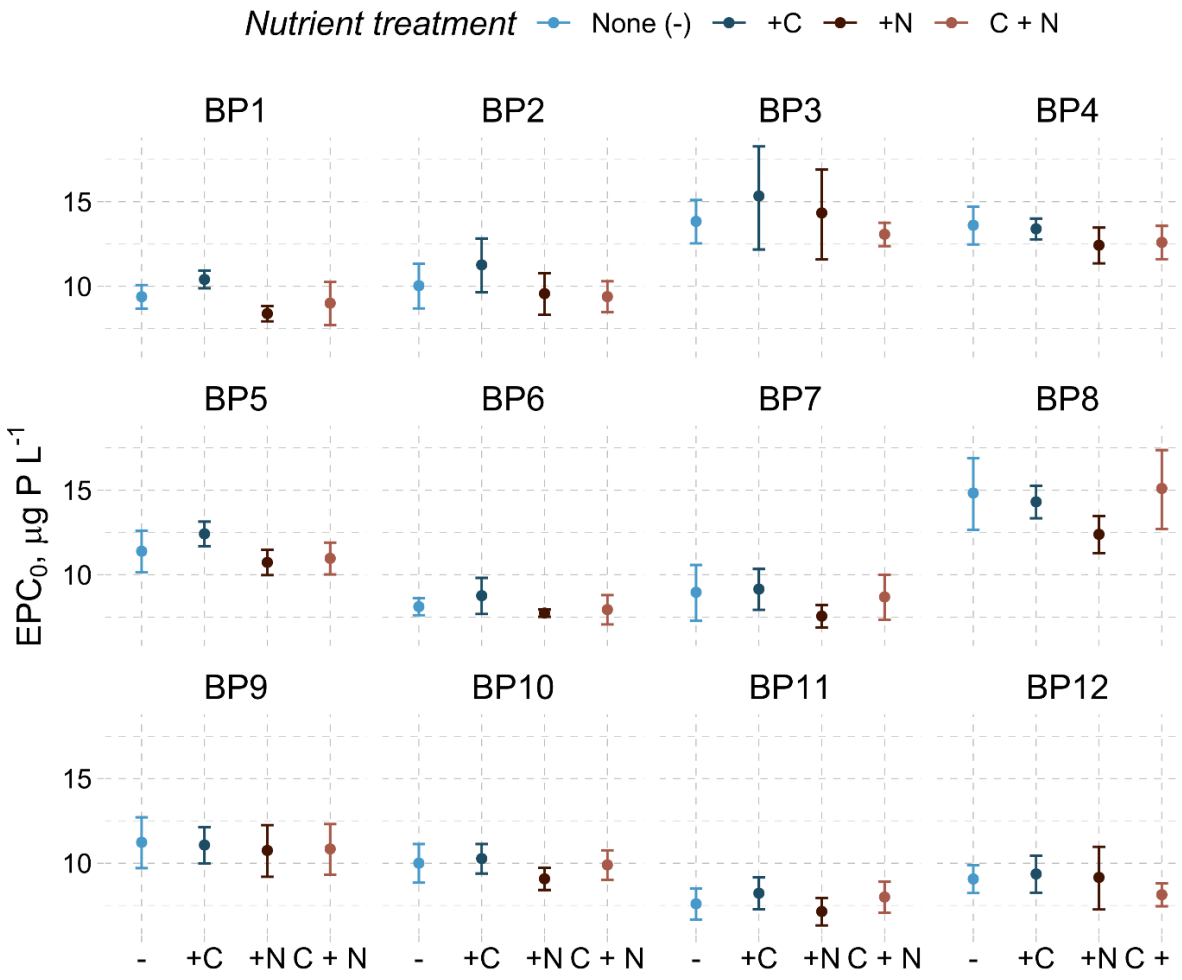

Figure 5.5 Sediment EPC_0 and its 95% confidence interval (via individual linear fits) for sediments from all 12 study streams and for four nutrient treatments: none (i.e., the original EPC_0 method, denoted here as -), +C (C:P of 200:1), +N (N:P of 20:1), and C+N (C:N:P of 200:20:1). Note that γ-irradiated sediments were also analyzed for the C+N treatment but this data did not differ from the original EPC_0 data.

+ N; p = 0.62) on EPC_0 did not out-perform the simpler model with only the +N effect. Though statistically significant, the +N effect on EPC_0 was relatively small, with a change in EPC_0 of -0.56 (95% C.I.; -0.81 to -0.31) µg P L^{-1}. Higher order effects (e.g., +N effect on EPC_0 varying per individual sediment) were not warranted by the data, neither were nutrient effects on sorption slopes.

5.5 Discussion

5.5.1 Microbial biomass as a stock of phosphorus: lysed P accounts for changes in EPC_0

Upon γ-irradiation, the flush of P from the microbial biomass (median of 10 mg P kg^{-1}) was free to adsorb onto the sediment, increasing EPC_0 (Figure 5.2 and Figure 5.3). While this result does not speak to the direct comparison of biotic and abiotic contributions towards sediment P uptake, it does show the potential influence of the sediment microbial biomass as a stock of P in streams. Similar to their contribution towards total P cycling in terrestrial ecosystems (Cleveland and Liptzin 2007; Turner et al.

Table 5.3 Model summary for the nonlinear mixed effects fit for the nutrient amendment experiment. Fixed effects include the two parameters, EPC_0 (*x*-intercept) and β (slope), plus a term for the +N effect on EPC_0 – other nutrient effects were not significant. Random effects include deviations in EPC_0 per sediment and the remaining within-group error. C.I. is confidence interval.

Term	Units	Estimate	95% C.I.
Fixed effects			
EPC_0	µg P L^{-1}	10.8	9.58 – 12.0
N effect on EPC_0	µg P L^{-1}	-0.560	-0.808 – -0.312
β	mg P kg^{-1} per µg P L^{-1}	0.01982	0.01976 – 0.01989
Random effects			
EPC_0 by sediment	µg P L^{-1}	2.15	1.43 – 3.22
Within sediment group error	mg P kg^{-1}	0.0259	0.0242 – 0.0277

2013), the microbial biomass represents a significant fraction of the total P in stream ecosystems. Importantly, sediment microbial biomass P is likely a very transient P pool, therefore contributing to P dynamics throughout the stream network (Mulholland et al. 1997; McDowell 2003; McDowell and Sharpley 2003b; Cross et al. 2005).

Few data exist on microbial biomass P in stream and river sediments, with ranges (using the same method from Brookes et al. (1982)) varying from: 0.5 to 10 mg P kg^{-1} in a 3rd to 4th order catchment having intensive forestry and pastoral land use (McDowell 2003), 2.4 to 23 mg P kg^{-1} in the present study, and 10 to 45 mg P kg^{-1} in a large river subject to considerable agricultural and urban pollution sources (Jaiswal and Pandey 2019). Streams with less P inputs will likely have much less sediment microbial biomass P (e.g., site BP1 in this study) though this also depends partly on the sediment (see below). For comparison, soil microbial biomass P concentrations can be an order of magnitude greater, varying from 3 to 430 mg P kg^{-1} (Cleveland and Liptzin 2007), likely due to the relatively greater organic matter content (Sinsabaugh et al. 2012). Considering that biomass C:P ratios are often more P-rich in sediment microbes (about 60:1; Cleveland and Liptzin 2007; Sinsabaugh et al. 2009) compared to periphyton (roughly on the order of 10^2 to 10^3; Cross et al. 2005), future work could investigate sediment microbial biomass as an important and perhaps variable stock of P in streams.

We assumed all of sediment microbial biomass P would be adsorbed by sediment surfaces and at a linear rate. However, we urge caution in applying this assumption to other sediments, owing to potential nonlinear P sorption (Froelich 1988; House 2003) and that amounts of microbial biomass P likely differ. Nonetheless, the consistency with actual changes in EPC_0 following γ-irradiation corroborates the hypothesis that the microbial biomass P stock can re-adsorb onto sediment surfaces, thereby increasing EPC_0 following sterilization. Since γ-irradiation is the least likely among sterilization methods to alter sediment physicochemical properties (Eno and Popenoe 1964; Berns et al. 2008; Buessecker et al. 2019), we can reasonably assert that the re-distribution of P – and the resulting increase in EPC_0 – was primarily due to cell lysis (Meyer 1979; Klotz 1985).

5.5.2 Reconciling past experiments on biotic vs. abiotic sediment P uptake

We suspect that our results may explain two of the inconsistences in past attempts to separate biotic and abiotic P uptake in sediments: sterilization techniques and a lack of microbial biomass P data. Firstly, more disruptive sterilization techniques than γ-irradiation have generally been used, namely, autoclaving and biocides. For example, when analyzing autoclaved sediments, several studies observed impressive differences between P sorption metrics for fresh and autoclaved sediments, concluding that up to ~40% of the total sediment sorption capacity was biotic (Haggard et al. 1999; Ryan et al. 2007; McDaniel et al. 2009). However, autoclaving likely changed the physicochemistry of these sediments since, in addition to cell lysis, autoclaving can disrupt sediment organic matter, alter surface chemistry, and increase some extractable elements such as Al, Fe and Mn (Meyer 1979; Wolf and Skipper 1994; Trevors 1996; Buessecker et al. 2019). In two comparison studies, Meyer (1979) and Klotz (1985) assessed potential biotic influence on sediment P uptake by comparing sediments analyzed fresh, autoclaved, and either γ-irradiated or with a phosphorylation inhibitor (cabonyl-cyanide *m*-chlorophenylhydrazone; CCCP), respectively. They found that autoclaving sediments released more P but also changed P sorption kinetics and capacity. In contrast, γ-irradiated sediments had sorption kinetics comparable to that of fresh sediments but had an increase in EPC_0 (attributed to cell lysis), while CCCP-treated sediments had no change in sorption capacity relative to fresh sediments. Similarly, Triska et al. (2006) found autoclaved sediments released more than 10× more P than fresh sediments in desorption assays and retained ~40% less P in sorption assays whereas a CCCP-treated sediment was comparable to fresh sediment for both sorption and desorption.

Aside from autoclaving, various biocides (e.g., $HgCl_2$, azide) have been applied to separate biotic and abiotic P uptake in stream sediments. Results with biocides have ranged from small increases in EPC_0 (2 µg P L^{-1}; Griffiths and Johnson 2018) to large or unrealistic changes in EPC_0 and sorption capacities (Lottig and Stanley 2007; Stutter et al. 2010). However, little is known about the potential unintended chemical interactions these biocides have with sediment surfaces. For example, $HgCl_2$ can react with organic matter, altering solution pH in the process (Buessecker et al. 2019), as noted by Griffiths and Johnson (2018) in their fresh vs. sterilized EPC_0 comparison. Buessecker et al. (2019) compared multiple methods to remove only the biotic mechanisms for soil N_2O production and concluded γ-irradiation to be the most suitable, as other methods (autoclaving, azide, $HgCl_2$, Zn, and chloroform) produced more physicochemical changes relative to fresh soil. We argue that such effects also apply to stream sediments. Neither autoclaving nor biocides will likely be accurate in removing solely the biotic reactions responsible for sediment P uptake. Thus, while not always acknowledged, conclusions for biotic and abiotic sediment P uptake reliant on autoclaving or biocides are most likely inaccurate.

In addition to disruptive sterilization methods, a second source for error in past estimates of biotic contributions to sediment P uptake was a lack of microbial biomass P data. As discussed, the microbial biomass can release a substantial flush of P upon lysis (Figure 5.3a) that can shift the equilibrium P concentration considerably. If our alternative hypothesis is correct – that the increase in EPC_0 with non-

disruptive sterilization is due to re-adsorption of the microbial biomass P following cell lysis (Figure 5.3b) – then we predict that the effect of sterilization would be even more pronounced for sediments with comparable microbial biomass P but weaker sorption capacities. Owing to their volcanic geology and high metal-oxide content (Dahlgren et al. 2004), our study sediments were relatively reactive and had high P sorption capacities (mean ASC of 54%); for comparison, the mean ASC from a national survey of sediments from 76 large New Zealand rivers was 9% (McDowell 2015). Interestingly, Munn and Meyer (1990) compared fresh and γ-irradiated sediment EPC_0 for two streams: one from a volcanic geology and one from a weathered granite geology. While both sediments had similarly low EPC_0 (5.4 and 1.0 µg P L^{-1}, respectively), the increases post-sterilization were respectively 3.2 and 9.7 µg P L^{-1}, consistent with a greater P sorption capacity in the volcanic geology. This increase was attributed to a predominance in biotic P uptake with no acknowledgement of possible bias due to the microbial biomass P.

5.5.3 Does carbon and nitrogen addition promote biotic P uptake and lower EPC_0?

Microorganisms in sediments demand nutrients according to the elemental stoichiometry of their biomass (Sterner and Elser 2002; Sinsabaugh et al. 2009). In contrast to lotic biofilm communities towards the top of the benthic zone (i.e., where light promotes autotrophic growth; Battin et al. 2016), and unlike for soil, sediment microbial communities (primarily heterotrophic) are more often C- or N-limited rather than P-limited (Hill et al. 2010, 2012; Sinsabaugh et al. 2012). Hence, we would only expect significant sediment biotic P uptake when C- and N-limitation is relieved. For sediments, microbial P-limitation is likely when C:P (molar basis) exceeds 60:1 and when N:P exceeds 7:1 (Cleveland and Liptzin 2007; Sinsabaugh et al. 2012). Under P-limitation, we expect that sediment microbes would sequester P for growth, thus potentially lowering the sediment EPC_0 through desorption. In other words, under P-limitation, sediment microbes may deplete P concentrations in solution below the sediment EPC_0, promoting P desorption from sediment surfaces and potentially lowering sediment EPC_0.

In this study, EPC_0 did not respond to C addition but decreased with N addition in both the +N and C+N treatments (-0.5 µg P L^{-1}; Figure 5.4 and Table 5.3). The +N effect here may have been dampened due to the large P (de)sorption capacity of these sediments (Small et al. 2016; Griffiths and Johnson 2018), as sediments display the greatest P buffering when near the EPC_0 (Froelich 1988). Indeed, the literature notes varied responses in P uptake to N supply. Using sediments in flume experiments at natural stream temperature, McDowell et al. (2017) observed relatively greater DRP uptake with N additions, a likely biotic effect since there was also a consistent increase in sediment microbial biomass P for the treatments with additional N. In contrast, Griffiths and Johnson (2018) measured stream P uptake rates while varying the background NO_3-N concentration: despite evidence for strong N- and P-colimitation, in-stream P uptake was unaffected by background N concentration, suggesting that stream P uptake was predominantly controlled by abiotic sediment P sorption rather than biotic P uptake. In a stream with plentiful N supply (mean of 1.1 mg NO_3-N L^{-1}), Oviedo-Vargas et al. (2013) measured P spiraling and

sediment EPC_0 during an experimental amendment of labile C (acetate) ~1 mg C L^{-1} above ambient conditions to stimulate sediment heterotrophic P demand. Despite greater C uptake during the experiment, P uptake was insensitive to the C addition (as whole-stream P uptake velocities and as sediment EPC_0). In their study, Oviedo-Vargas et al. (2013) suggested that the predominantly heterotrophic microorganisms within the sediments were meeting their P demands through organic P mineralization or from sediment P desorption rather than from the water column. Similarly, the sediments in the present study buffered solution P towards the EPC_0 through abiotic sorption processes in spite of biotic P uptake.

A reason for the response to N but not C may lie in the fact that labile C can increase microbial respiration (Roberts et al. 2007; Oviedo-Vargas et al. 2013; Demars 2019) but will not directly affect P concentrations. It is microbial production, not respiration, that carries a P demand (Sterner and Elser 2002). Sediment microorganisms here may have had enough C supply available to meet their demands for production, thereby not influencing EPC_0 for the +C treatment. For example, sediment biofilms can hold over C in their extracellular polymeric substances (Battin et al. 2016) or enzymatically degrade organic substrates (Findlay et al. 2003; Sinsabaugh et al. 2009; Kaplan and Cory 2016) to meet C requirements. Most of the catchments in our study had some amount of pasture (35 to 82% for sites BP2 through BP12), which may provide an ample source of C when soil pore waters hydrologically connect to the streams (Kaplan and Cory 2016; McNally et al. 2017). Additionally, these pasture-dominated catchments had diminished stream canopy cover, meaning likely greater autochthonous C supplies for the stream ecosystem due to increased primary productivity (Dodds 2007; Finlay 2011). This may have supplied sufficient high-quality C resources (e.g., from algal biomass turnover; Meyer 1994; Kaplan and Cory 2016) to the sediments to prevent any C-limitation before EPC_0 measurements.

Our results from labile C and N additions to EPC_0 measurements suggest little to no influence of biotic P demand on EPC_0 for these streams, with only a small overall effect arising from the removal of potential N-limitation (Table 5.3). Coupled with the results of the first experiment – that much of the apparent differences in P sorption between fresh and sterilized sediments is due to the microbial biomass P flush – this experiment suggests that sediment biotic P demand may not be a significant part of the EPC_0 as it is typically measured, i.e., with only P added. Rather, the benthic sediment EPC_0 may primarily reflect the sediment's abiotic sorption characteristics and prior P loading to that sediment (Froelich 1988; McDaniel et al. 2009; McDowell 2015). Our results come from streams that are likely N-limited and that drain a naturally P-rich geology. For other streams, however, if sediments are poorly sorptive (e.g., little Al/Fe oxides or coarse particle sizes) and P availability is low, then biota may have a relatively stronger influence on P uptake in sediments (Lottig and Stanley 2007). The dynamic between P sorbed to sediments and P immobilized in microbial biomass should be explored in other systems that vary in stream geomorphology, sediment characteristics, and nutrient inputs.

5.5.4 Improving our understanding of the role of microbes in stream P attenuation

The results from this study help to separate some of the processes usually lumped together in sediment P sorption and, at the greater scale, P attenuation in streams. Characterizing these processes more clearly will help target future research on sediment P reactions, for example, in determining the role of microbial biomass as a stock of P and its relationship with the abiotic sediment P compartment. When one measures EPC_0, the EPC_0 is unlikely to be affected by biotic uptake partly since P is often not the limiting nutrient for sediment heterotrophic microbes. Rather, EPC_0 is largely a function of the sediment reactive surfaces and the previous exposure of P to that sediment. Integrating this information and more biota-specific P reactions with hydrological fluxes will improve our understanding of stream and catchment-scale P biogeochemistry (Manzoni and Porporato 2011).

Sediment EPC_0 provides information on how sediments subsidize available P for microbial growth and so influence stream P biogeochemistry. Sediments may either maintain or accumulate more sorbed P ($DRP \geq EPC_0$) but, when local DRP concentrations are depleted through *inter alia* microbial P uptake ($DRP < EPC_0$), sediments may desorb P so nearby microorganisms can meet their stoichiometric constraints. If sediments drip feed P – according to the EPC_0 – to 'P-hungry' sediment biofilms (particularly heterotrophic bacteria), then we could expect a positive relationship between EPC_0 and microbial biomass P. Indeed, the two variables had a strong positive correlation (ρ=0.87, p=3.1e-4), suggesting that sediment microorganisms play a role in regulating P desorption from sediments back into the water column. This coincides with what we know on the coupled biogeochemical cycling of C, N, and P in streams (Oviedo-Vargas et al. 2013; Maranger et al. 2018) and has important consequences. For example, the P subsidy provided by stream sediments (as approximated by EPC_0) may promote greater C and N processing in streams (Stelzer et al. 2003; Findlay and Sinsabaugh 2006; Tank et al. 2010). Another consequence, however, is that this P supply could support a large sediment biofilm, meaning the biofilm may clog hyporheic flowpaths (Hartwig and Borchardt 2015; Battin et al. 2016) thus limiting the interaction between reactive sediment surfaces and P in the water-column (Boano et al. 2014; Weigelhofer et al. 2018a). Further exploring the sediment P buffer will likely aid us in better characterizing stream P biogeochemistry and its consequences (Hamilton 2012).

Overall, considering the dynamic between the sediments and their microbial biofilms will help in understanding the stream sediment P buffer and hence stream P attenuation. This understanding can sharpen our view of stream P biogeochemistry but also help in understanding the effects of our efforts in mitigating nutrient pollution, e.g., from changing wastewater inputs (Scott et al. 2011; Wilcock et al. 2020) or from management on-land (Meals et al. 2010; McDowell et al. 2018). This recognition of both biotic and abiotic mechanisms can help avoid unrealistic expectations by policy agencies in the speed or efficacy of such efforts to mitigate P pollution.

Chapter 6

Distinguishing phosphorus uptake by periphyton and sediment in an open-canopy stream

6.1 Abstract

Several processes influence phosphorus (P) availability in streams at baseflow. Periphyton P uptake is presumably a function of gross primary productivity (GPP) and can be a substantial part of P removal in many streams. However, benthic sediments also buffer P concentrations in solution through P sorption, which is often overlooked in studies of stream P cycling. To separate these two processes, we conducted P tracer pulse-injection experiments in a 4th order, open-canopy stream during dark ('dawn') and light ('noon') conditions. Stream periphyton contributed a high GPP during the day (about 5 g O_2 m^{-2} d^{-1}) which was matched by the stream ecosystem respiration (ER; about -5 g O_2 m^{-2} d^{-1}). Cumulative GPP rates during the first 2.5 h of the dawn and noon injections were, respectively, 0.13 and 1.18 g O_2 m^{-2} – an order of magnitude difference. However, whole-stream ambient uptake rates of dissolved reactive P (DRP) were near equal: 0.39 to 0.61 µg P m^{-2} s^{-1} and 0.37 to 0.41 µg P m^{-2} s^{-1}, respectively. Benthic sediments were strongly sorptive for P, with fast P sorption rates (majority of sediment P sorption in incubation experiments occurred within <10 min) and an equilibrium phosphate concentration at net-zero sorption (EPC_0) of about 8 µg P L^{-1}, indicating a potential for P removal throughout the experiment (ambient DRP was 15 µg P L^{-1}). Taken together, these results suggest that periphyton had relatively little impact on P uptake in this stream while sediment P sorption had sufficient capacity for P removal during the tracer passage in both injections. This work highlights the sediment P buffer as a prominent but previously under-appreciated mechanism for P attenuation in streams.

[SCENE: The student and his advisors plan the stream P cycling experiment in a whirlwind of a meeting. The discourse currently swirls around periphyton and its role in P cycling in the stream.]

LEO [perturbed, miffed]:

What even is periphyton?!

RICH [tepid, wry]:

It's grass, mate.

6.2 Introduction

Multiple processes determine the availability of phosphorus (P) in streams. Streams can strongly remove P at baseflow, as demonstrated by many experimental in-stream studies (Ensign and Doyle 2006; Haggard and Sharpley 2007; Weigelhofer 2017). Yet, these experiments often quantify bulk, rather than process-specific, uptake parameters and so our ability to predict stream P uptake and longer-term P retention under different conditions remains limited. Parameterizing the processes responsible for P uptake at varying locations, times, and scales throughout the stream network is key for the prediction of watershed P transport (Hamilton 2012; Dupas et al. 2019; Frei et al. 2020).

In a select few cases, enough information is available to reliably attribute P retention to one or more specific processes. Pioneering work with $^{33}P/^{32}P$ tracers measured or modeled the fluxes of P into various biotic stocks (periphyton, detritus, microbial biomass and more) in small forested streams (Newbold et al. 1983; Mulholland et al. 1983, 1994, 1997; Elwood et al. 1988). These studies emphasized benthic periphyton as a major but changeable biotic P store. Large-scale patterns in periphyton and P data also highlight the importance of this biotic stock in P retention (Biggs 2000; Cross et al. 2005; Dodds 2007). Additionally, microbes within the sediment (i.e., the biofilm beneath the substrate exposed to light; Battin et al. 2016) may demand P, and certainly are an overlooked P stock in streams, though these biota are also commonly limited by carbon (C) and/or nitrogen (N) (Sinsabaugh et al. 2012).

As for abiotic processes in streams, benthic sediments constitute another major but changeable P store via P sorption (Hamilton 2012; Simpson et al. 2021). For example, after experimentally enriching a tropical stream with P for 8 years, Small et al. (2016) estimated that ~99% of excess P in the stream was sorbed to sediments, while organic matter stores of P were minimal. Stream benthic sediments can strongly adsorb P depending on texture, geochemical characteristics, and whether P concentrations are above an equilibrium point (known as EPC_0; if below this concentration, sediments may desorb P back to the water-column; Froelich 1988; Simpson et al. 2021). Additional abiotic processes will largely depend on the stream/catchment characteristics, such as catchment geology for calcite (co-)precipitation of phosphate (Jarvie et al. 2006a; Corman et al. 2016).

Both biotic and abiotic processes contribute to in-stream P retention, and occur simultaneously (Stutter et al. 2010; Weigelhofer et al. 2018a), making stream P retention difficult to predict. However, a fruitful strategy may be to leverage the fact that drivers of varying P retention processes are biogeochemically distinct. As an example, for P uptake via primary producers (i.e., periphyton for most shallow streams), there needs to be sufficient light, temperature, nitrogen, and stable, non-scouring flows (Biggs 2000; Dodds 2003; Biggs et al. 2005; Fanta et al. 2010). More succinctly, assuming negligible luxury P uptake, periphyton P demand is primarily a function of gross primary productivity (GPP; Mulholland et al. 1994; Mulholland 1996). Stream GPP may be paired with other measurements to infer how photoautotrophs control nutrient availability (Hensley and Cohen 2016; Jarvie et al. 2018). Recent advances in monitoring dissolved oxygen and modelling allow the estimation of whole-stream ecosystem

metabolism. In contrast, while sediment P sorption can vary with streamwater conditions such as temperature and pH, it is primarily a function of sediment sorption affinity, past P loading, and the current P concentration (Simpson et al. 2021). Therefore, within a daily cycle for a stream at baseflow, sediment P uptake may be more static relative to periphyton P uptake (Cohen et al. 2013; Martí et al. 2020), meaning that diel variation in P uptake may identify process-specific P retention in streams.

Considering that (1) likely P retention processes can be narrowed for some streams to periphyton P uptake and sediment P sorption, (2) stream periphyton likely demand P most strongly when GPP occurs, and (3) periphyton GPP requires light, 'light vs. dark' experiments may be a way to distinguish periphyton P uptake and sediment P sorption in streams. Here, we report a P tracer experiment conducted under light and dark conditions in an open canopy, 4th order stream. We hypothesized that sediment P uptake would be equivalent for similar hydrologic conditions within a short time window (several days of stable weather), but periphyton P uptake would vary diurnally as a function of GPP. Previous work at the study stream established that calcite precipitation was unlikely even at peak periods of GPP and that benthic, fine sediment microbes (largely heterotrophs) had negligible P demand compared to sediment P sorption (Simpson et al. 2020). Therefore, by performing pulse injections of P at solar dawn and noon in an open-canopy stream ('dawn' and 'noon' injections, respectively), we experimentally estimate the contributions of periphyton and sediments to whole-stream P uptake.

6.3 Methods

6.3.1 Study stream site

The study stream was the Kaituna River, located on Banks Peninsula, Canterbury, New Zealand. The Kaituna River typifies many streams on the peninsula, and drains a mixed land use catchment, progressing towards predominantly pastoral land-use towards the outlet at Te Waihora/Lake Ellesmere. The 33.1 km^2 catchment is 70% grassland (pastoral land use) with some forestry (19%) and shrub/scrub (11%) situated in the upper hills. Notable characteristics of the stream are its volcanic-basic geology (Snelder and Biggs 2002), riffle-pool morphology, and its open canopy in the lower part of the valley. The surrounding hills provide some topographic shading in the early and late parts of the day, leading to sharp changes in light availability to the stream. The study reach on Kaituna River has a slope of 0.0211 m m^{-1} and mostly gravel to gravel/cobble substrate. Macrophytes were scarce in the study reach at the time of the experiment.

6.3.2 Hydrology

Hydrological characteristics of the study stream are summarized in Table 6.1. We measured stream discharge (Q, m^3 s^{-1}) via the salt slug method (Rantz 1982; Baker and Webster 2017). We measured discharge near the time of tracer injections; however, due to instrument failure, we did not measure discharge on the day of the dawn injection (below) and instead use data collected five days later to interpolate discharge for the dawn injection. A stream gage ~3 km downstream, maintained by

Environment Canterbury, suggested little change in baseflow during this period. While no surface lateral inflows were present during the study, we estimated a lateral groundwater inflow (Q_L) of 0.0545 L s^{-1} m^{-1} by measuring discharge at two locations (110 m apart) and assuming linear changes in discharge with longitudinal length. No rainfall was recorded around the January 2020 injections.

Wetted widths and mean depths were measured at 10 transects uniformly spaced along the stream length. Longitudinal stream distances between the injection point (for salt slugs and the tracer slugs) and monitoring sites (reach lengths; see below) were determined with Google Earth via the measure tool. The accuracy of this tool was verified by measuring easily identifiable structures at the site (e.g., the edges of a pump house); digital map measurements were within ~10 cm of field measurements, and hence appropriate for the reach scale (100s of m).

6.3.3 Tracer injections and water sampling

Our study design consisted of two slug tracer injections at the Kaituna study reach: one near solar noon (14:00; solar noon was at 13:39) on January 17th, 2020 and one beginning just before dawn (06:00; sunrise was at 06:13) on January 21st, 2020. Henceforth, these will be referred to as the 'noon' and 'dawn' injections, respectively. By completing both injections in a narrow time window, we minimized possible differences in stream hydrology, periphyton biomass, and sediment characteristics between injections.

All injections started upstream of a riffle in an area of concentrated, turbulent flow to ensure good mixing. Sampling stations were located 135 m (B1), 237 m (B4), and 338 m (B5) downstream of the injections. Each tracer slug consisted of 750 g NaBr and 250 g KH_2PO_4 (i.e., 582.4 g Br and 56.9 g P) fully dissolved in a 2 L bottle of deionized water which was further mixed with stream water in a 10 L bucket prior to injection; we targeted enough P mass to observe biotic P uptake at P saturation for at least several min of the slug passage (Covino et al. 2010). Additionally, we targeted a high Br concentration to improve sensitivity in the receding tail of the breakthrough curve (i.e., maintain concentrations well above detection limits and background variability) and therefore improve estimation of solute transport in this critical region of the breakthrough curve (Drummond et al. 2012). Grab samples were collected from the centroid of flow at all three stations with minimal disturbance of the

Table 6.1 Hydrology of study reach at Kaituna River. Discharge and nominal travel time are for site B4 (237 m from injection point).

	Units	Noon	Dawn
Slope	m m^{-1}	0.0211	
Mean wetted width	m	5.3	5.3
Mean depth	m	0.16	0.16
Mean cross-section area	m^2	0.64	0.64
Discharge	m^3 s^{-1}	0.093	0.088[a]
Mean velocity	m s^{-1}	0.113	0.105
Nominal travel time	min	35	40

[a]Because of instrument failure, a discharge measurement was taken 5 days later after the dawn injection (81 L s^{-1}) and this value is given assuming a linear change over time.

stream. These samples were immediately filtered (0.45 µm) and placed on ice in the dark until later being refrigerated in the laboratory. Nominal sampling frequencies ranged from 30 seconds (near peak), 1-5 min (pre- and post-peak), and 5 to 30 min for the tail based on prior data from the discharge salt slugs. We targeted ~60 samples per site for an injection with more samples concentrated near critical points in the breakthrough curve, e.g., when changes in concentration over time were greatest (Wagner and Harvey 1997; Covino et al. 2010).

6.3.4 Water analyses

Dissolved reactive P (DRP) was measured on filtered samples within 24 h. We used the molybdenum-blue method (Murphy and Riley 1962) with a 5 cm quartz cell for concentrations below ~160 µg P L^{-1} (method detection limit of ~2 µg P L^{-1}) and a 1 cm cell for higher concentrations. Quality control checks and occasional duplicate analyses established a relative error of 5%. Anions (Br, Cl, NO_3, and SO_4) were measured on filtered samples via ion chromatography (detection limits of 0.02, 0.012, 0.011, and 0.50 mg L^{-1}, respectively). On select filtered stream samples, we measured dissolved cations (ICP-OES), dissolved organic C (TOC analyzer), mineral N (flow-injection analysis), and alkalinity (Rounds 2012). On selected unfiltered stream samples reflecting ambient conditions, we measured total suspended solids (method 2540 D; APHA 2005) and total P (method 4500; APHA 2005).

6.3.5 Periphyton sampling and analyses

Following the dawn injection, benthic periphyton were sampled at 10 uniformly spaced transects along the study reach following Biggs and Kilroy (2000). Between 100 and 200 cm^2 of surface area was sampled per transect by uniformly sampling gravel/cobble (~10 stones). A known surface area of periphyton was scrubbed off with a toothbrush and rinsed into a container. The slurries were kept in the dark on ice and later frozen (-20 °C) in the laboratory until analysis.

After thawing, blending the slurries, and making up to a standard volume with deionized water, one well-mixed aliquot was filtered through a glass fiber filter (previously combusted at 400 °C for 1 h): filter weights were recorded following 24 h at 105 °C and after 4 h at 400 °C to measure dry mass and ash-free dry mass (AFDM), respectively. Another slurry aliquot was filtered and the filters extracted for chlorophyll *a* (Chl *a*) in ethanol (96%; incubated 5 min at 80 °C followed by 24 h at -20 °C) before Chl *a* concentrations were analyzed via spectrophotometry following Parker et al. (2016). During the acidification step to correct absorbances for phaeopigments (re-measuring absorbances after conversion of Chl *a* to pheophytin *a*), we allowed ~45 min (but <60 min) of reaction time to get stable conversion of extracts (Parker et al. 2016). All sample handling for Chl *a* analyses were carried out under minimal light conditions.

Leftover periphyton slurry was frozen and lyophilized. Freeze-dried periphyton mass was analyzed for total C, N, and P content via dry combustion (C and N; Rutherford, et al. 2008) and via ICP-OES following a digestion (P; Miller 1997).

6.3.6 Sediment analyses

Following periphyton sampling, benthic sediment (top 1-4 cm) was collected with a shovel between stations B4 and B5. We targeted advective zones over a ~10 m stream length. We sieved the material to <2 mm, allowed the composite sample to settle for 30 min before decanting excess water, and stored the sample cool and in the dark until refrigerated (4 °C) in the laboratory.

A subsample of sediment was oven-dried (105 °C) overnight to measure moisture content; analyses on fresh (wet) sediments below are on a dry weight (d.w.) basis. Another subsample was freeze-dried to later analyze for total C and N concentrations (via C/N Elemental Analyzer; Carter and Gregorich 2007) and for particles sizes <1 mm via laser diffraction (Eshel et al. 2004). Sediment pH was measured on fresh sediments in deionized water with a 1:5 g g^{-1} (sediment:solution) after 30 min equilibration. As a measure of sediment microbial respiratory activity, we measured dehydrogenase activity on fresh sediments with a triphenyltetrazolium chloride substrate (Öhlinger and Von Mersi 1996).

Water extractable P was measured on fresh sediments by shaking sediments with deionized water (1:10 g g^{-1}) for one hour, centrifuging, and analyzing for DRP. Anion storage capacity (ASC, in %), a measure of potential P retention, was measured following Saunders (1965). Equilibrium phosphate concentration at net zero sorption (EPC_0) was measured on fresh sediments as described by Simpson et al. (2019). With the incubations for EPC_0, we also varied incubation times to examine sorption kinetics, with times of 10, 30, 60, 240, and 960 min; we refer to the longest time (960 min) as most closely resembling 'equilibrium'.

6.3.7 Stream ecosystem metabolism

We collected data for estimating whole-stream metabolism parameters via the one-station method (Hall and Hotchkiss 2017). At least one day before tracer injections, we positioned two light loggers (Onset HOBO MX2202) at the downstream end of the study reach: one on the streambank to measure top-of-stream irradiance and one submerged onto the streambed near the thalweg to measure in-stream irradiance. Both light loggers recorded temperature and light intensity (lux) at 1 min intervals. These light intensity data were later converted to photosynthetic active radiation (PAR; µmol photons m^{-2} s^{-1}) via the 'sand' calibration given by Long et al. (2012). Before matching up to dissolved oxygen (DO) timeseries (below, 5 min frequency), the PAR data were smoothed to reduce noise and improve metabolism parameter estimation. For this, we used a generalized additive model (Wood 2017) with cyclic smooths (diel signal), autoregressive model error (order = 1), and a thin-plate regression spline to pick up residual variability (e.g., due to cloud cover).

At the same location in-stream, a DO probe (YSI Professional Plus with a polarographic DO probe) was situated perpendicular to streamflow and recorded on 5 min intervals. As a backup later in the experiment, we also deployed a HACH HQ40D meter with an optical DO probe. Both DO probes were calibrated in air-saturated water by bubbling air into a bucket of stream water on-site (Hall and Hotchkiss

2017). Calibrations were checked daily and data were corrected, where possible, if instrumental drift occurred. Unfortunately, there were periods of instrument failure where no DO data were recorded (~12 h on January 19th and ~22 h on January 20th). Therefore, we focus on the first 3 days of DO data for estimating metabolism models and treat the later, isolated series of DO as sets to validate or test the models. We refer to these periods of stream metabolism data as 'calibration' and 'validation', respectively, when evaluating stream metabolism model performance.

Climatic data from a nearby (12 km) monitoring station (via CliFlo, station network number H32674; https://cliflo.niwa.co.nz/) provided surface irradiance (W m^{-2}) and mean sea level barometric pressure (hPa). Surface irradiance was converted to PAR via the factors compiled by Holtgrieve et al. (2010) to corroborate the benthic surface PAR data described above.

Stream metabolism was modelled from the DO timeseries via the following governing equation:

$$\frac{dO_2}{dt} = \frac{GPP}{z} + \frac{ER}{z} + K(O_2^{sat} - O_2)$$

where, O_2 is DO (mg O_2 L^{-1}), *t* is time (d), GPP is gross primary productivity (a positive flux; g O_2 m^{-2} d^{-1}), ER is ecosystem respiration (a negative flux; g O_2 m^{-2} d^{-1}), *z* is mean stream depth (m), *K* is the stream reaeration coefficient for O_2 (d^{-1}), and O_2^{sat} is the saturated DO concentration depending on surface pressure and stream temperature. Since lateral flow inputs were absent and groundwater flows were negligible along the stream reach, we did not account for these inputs for the DO flux (Hall and Tank 2005).

For the GPP term in the model, we tested a linear function of PAR (*I*):

$$GPP = \alpha I$$

and a saturating function of PAR (Jassby and Platt 1976):

$$GPP = P_{max} tanh\left(\frac{\alpha I}{P_{max}}\right)$$

where α is the slope of the light-saturation curve (g O_2 m^{-2} d^{-1} per unit PAR) and P_{max} is the GPP at optimal light intensity (g O_2 m^{-2} d^{-1}). The saturating function is preferable for periods of high light intensities as photosynthetic activity in some stream periphyton can begin to saturate (Boston and Hill 1991). We assumed the periphyton communities in this open-canopy stream did not experience photoinhibition since periphyton biomass was undisturbed (i.e., high accrual) and ambient DRP was relatively plentiful (>10 µg P L-1; Boston and Hill 1991; Hill et al. 2009).

For the stream ecosystem respiration (ER) term, we examined two assumptions: (1) ER was constant, reflecting a stable carbon (C) substrate supply for benthic communities throughout the day (Hall and

Hotchkiss 2017) and (2) ER varied with stream temperature (~9 °C fluctuation within some days). For the latter, we modeled ER as a function of stream temperature via Arrhenius kinetics (Acuña et al. 2008; Jankowski et al. 2014):

$$ER = ER_{ref} \times exp\left\{-E\left(\frac{1}{k_B T} - \frac{1}{k_B T_{ref}}\right)\right\}$$

where ER_{ref} is the ER at a reference temperature (T_{ref}, K), T is the stream temperature (K), E is the activation energy (eV), and k_B is the Boltzmann constant (8.617×10^{-5} eV K^{-1}). The temperature sensitivity of whole-stream ER is reflected in the E parameter. Here, we used a T_{ref} of 15 °C (288 K). For the alternative model of constant ER, E can be considered 0 and so the above expression simplifies to ER = ER_{ref}.

The last term in the metabolism model refers to the physical gas transfer between the water column and the atmosphere. For this, we modeled K_{600} as a parameter, which corresponds to the reaeration coefficient for a Schmidt number of 600 (Raymond et al. 2012). For the corresponding timestep and stream temperature, K_{600} was converted to the O_2-specific K before calculations (Hall and Hotchkiss 2017).

We tested metabolism models ranging in complexity with (1) GPP as either a linear or saturating function of PAR and (2) ER as either constant or temperature-sensitive, hence, four models. All models were fitted using maximum likelihood estimation via the 'maxLik' package (Henningsen and Toomet 2011). For each fit, we calculated the Nash-Sutcliffe efficiency, percent bias, and root mean square error (RMSE) as metrics to assess model performance (Moriasi et al. 2007). We compare these metrics for calibration, validation, and all data (see above). Further, to quantify model parameter uncertainty, we calculated bootstrap parameter distributions by randomly re-sampling model residuals and re-estimating parameters (sensu Bogert et al. 2007). We used 10,000 bootstrap samples and report the median and 95% confidence interval for each parameter.

6.3.8 In-stream phosphorus uptake: solute spiraling via TASCC

Using the slug injections of Br (the conservative tracer) and DRP (reactive tracer), we analyzed in-stream DRP uptake via the solute spiraling framework (sensu Stream Solute Workshop 1990). Specifically, we estimated DRP spiraling metrics from the slug injections via the Tracer Additions for Spiraling Curve Characterization (TASCC) method of Covino et al. (2010). Detailed steps behind the calculations are given by Covino et al. (2010) and Brooks et al. (2017).

Briefly, with Br as the conservative tracer concentration (*cons*) and DRP as the reactive nutrient tracer concentration (*nut*), both tracer concentrations were monitored ($cons_{tot\text{-}obs}$ and $nut_{tot\text{-}obs}$, respectively; all concentrations indicated here are considered as mg m^{-3}) and background-corrected ($cons_{add\text{-}obs}$ and $nut_{add\text{-}obs}$) by subtracting ambient concentrations (nut_{amb} and $cons_{amb}$) at each sampling station. For a given

station, each observation in the breakthrough curve was used to calculate a dynamic solute spiraling length of the added nutrient ($S_{w\text{-}add\text{-}dyn}$; m) via:

$$S_{w-add-dyn} = \frac{-L}{log\left(\frac{nut_{add-obs}}{cons_{add-obs}}\right) - log\left(\frac{nut_{inj}}{cons_{inj}}\right)}$$

where L (m) is the length of the reach from the injection point and nut_{inj} and $cons_{inj}$ are the *nut* and *cons* concentrations of the injectate solution. The dynamic uptake velocity ($v_{f\text{-}add\text{-}dyn}$; m s^{-1}) is calculated as:

$$v_{f-add-dyn} = \frac{Q}{S_{w-add-dyn} \times w}$$

where w is the stream wetted width (m). This may be re-expressed as dynamic areal nutrient uptake rate ($U_{add\text{-}dyn}$; mg m^{-2} s^{-1}):

$$U_{add-dyn} = v_{f-add-dyn} \times nut_{add-dyn}$$

where $nut_{add\text{-}dyn}$ is the dynamic concentration of the added nutrient is the following geometric mean:

$$nut_{add-dyn} = \sqrt{nut_{add-obs} \times nut_{cons}}$$

and nut_{cons} is the added nutrient concentration corrected for conservative transport:

$$nut_{cons} = cons_{add-obs} \times \frac{nut_{inj}}{cons_{inj}}$$

The total dynamic nutrient concentration ($nut_{tot\text{-}dyn}$), which accounts for ambient concentrations, is:

$$nut_{tot-dyn} = \sqrt{nut_{tot-obs} \times (nut_{cons} + nut_{amb})}$$

By regressing $S_{w\text{-}add\text{-}dyn}$ against $nut_{tot\text{-}dyn}$, the y-intercept corresponds to ambient nutrient uptake length, $S_{w\text{-}amb}$ (m). The associated ambient uptake metrics, $v_{f\text{-}amb}$ and U_{amb}, are calculated as in the above expressions but with substitution of $S_{w\text{-}amb}$ for $S_{w\text{-}add\text{-}dyn}$ and nut_{amb} for $nut_{add\text{-}dyn}$. Finally, the total dynamic areal uptake is calculated as:

$$U_{tot-dyn} = U_{add-dyn} + U_{amb}$$

For all these calculations in the TASCC method, we only analyzed the receding limb of the breakthrough-curves to focus on periods where the tracers had time to interact with the benthic and hyporheic zones (Brooks et al. 2017; Li et al. 2021). Further, while other studies have modeled $U_{tot\text{-}dyn}$ as a saturating function of $nut_{tot\text{-}dyn}$ (including for DRP) via Michaelis-Menten kinetics (Covino et al. 2010; Weigelhofer et al. 2018a; Griffiths and Johnson 2018), we found limited evidence for this in our data and therefore simply examine $U_{tot\text{-}dyn}$ as a linear or log-linear function of $nut_{tot\text{-}dyn}$; we test for

differences in this relationship between sites and between the dawn and noon injections in a linear modelling exercise.

6.4 Results

6.4.1 Stream, periphyton, and sediment characteristics

The study stream was at typical summer baseflow conditions with similar discharge (0.088 to 0.093 m^3 s^{-1}) for both injections (Table 6.1) and with no recent storm events. Stream pH varied by ~1 S.U. between injections (7.4 to 8.5), reflecting differences due to photosynthesis; potential for calcite precipitation at this pH was negligible considering the moderate alkalinity and Ca concentrations. Both NH_4 and NO_3 concentrations remained below detection limits during the injections (Table 6.2) while ambient DRP concentrations were ~15 µg P L^{-1}.

Under the warm, stable conditions prior to our study, the benthic substrate in the stream accumulated considerable periphyton mass, with average AFDM of 17.8 g m^{-2} and Chl *a* content of 50.6 mg m^{-2} (Table 6.4). Periphyton mats were filamentous (up to a few cm long) and easily sloughed off surfaces. Periphyton dry-mass elemental ratios (molar basis) averaged 9.8 C:N, 140 C:P, and 14 N:P.

Table 6.2 Stream water physicochemistry for noon and dawn tracer injections. These are a mixture of continuous monitoring and grab samples: specific conductivity through PAR are reported as means during the injection; alkalinity through NO_3-N and TP are from select grab samples; DRP is the ambient DRP concentration at site B4 at beginning of the tracer injection. Some analytes were not measured for the noon injection. BD refers to below detection, blank cells refer to no measurement.

	Units	Noon	Dawn
Specific conductivity	µS cm^{-1}	166	176
pH	S.U.	8.50	7.37
Temperature	°C	21.6	17.5
DO saturation	%	124	85
PAR at benthic surface	µmol m^{-2} s^{-1}	2015	135
Alkalinity	mg L^{-1} as $CaCO_3$		53.3
Ca	mg L^{-1}		11.5
Fe	mg L^{-1}		0.24
TSS	mg L^{-1}		1.8
DOC	mg L^{-1}		1.84
NH_4-N	mg L^{-1}	BD	BD
NO_3-N	mg L^{-1}	BD	BD
DRP	µg P L^{-1}	15	14
TP	µg P L^{-1}		30

Table 6.4 Stream benthic periphyton measurements for the January 2020 injections (both noon and dawn). All values are the mean (SD) from 10 transects. Molar ratio SD's were calculated considering the component uncertainties for both elements (Sterner and Elser 2002).

	Units	**Mean (SD)**
Ash free dry mass (AFDM)	g m^{-2}	17.8 (7.8)
Chl *a* content	mg m^{-2}	50.6 (41)
Autotrophic index[a]	g AFDM g^{-1} Chl *a*	534 (302)
C content	g C kg^{-1} DM	106 (20)
N content	g N kg^{-1} DM	12.6 (2.0)
P content	g P kg^{-1} DM	1.96 (0.28)
C:N	mol C mol^{-1} N	9.81 (2.4)
C:P	mol C mol^{-1} P	140 (33)
N:P	mol N mol^{-1} P	14.2 (3.1)

[a]Based on Biggs and Kilroy (2000)

Benthic sediments (<2 mm) below the gravel/cobble substrate were relatively sandy (85% sand) yet moderately sorptive, with an ASC of 51 out of 100% (Table 6.3). As expected for sediments from this volcanic geology, sediment total P was relatively high (2400 mg P kg^{-1}). Sediment EPC_0, using a 16 h incubation time, was 7.66 (95% CI; 6.20 to 9.07) µg P L^{-1} (Figure 6.1). Notably, this sediment showed a high sorption capacity and rapidly sorbed P, producing nearly equivalent EPC_0 values when initial concentrations were below 100 µg P L^{-1}. EPC_0 calculations for each incubation time indicated a rapid convergence to the long-term (16 h) EPC_0; even the shortest incubation time (10 min) had an EPC_0 of 3.39 (1.67 to 5.05) µg P L^{-1}. Considering that ambient DRP in the stream during both injections (~15 µg P L^{-1}) was above the EPC_0, adsorption was likely to prevail over desorption throughout the study. Further, the linear relation between concentration and sorption across the wide range of DRP concentrations and time scales supports our use of simple 1st-order P uptake in this stream (as in the TASCC method, below).

Table 6.3 Benthic stream sediment physicochemical characteristics representative of both tracer injections (noon and dawn). Particle size data are based on particles below 1 mm diameter due to instrument limitation.

	Units	**Value**
pH	S.U.	7.35
particle sizes	clay / silt / sand, % vol.	1.3 / 13.4 / 85.3
Median particle diameter	µm	684
TC	g kg^{-1}	8.4
TN	g kg^{-1}	0.9
TP	mg kg^{-1}	2394
Anion storage capacity	%	50.7
EPC_0	µg P L^{-1}	7.7
Dehydrogenase activity	mg TPF kg^{-1} h^{-1}	2.13

6.4.2 Stream ecosystem metabolism

Table 6.5 reports the parameter estimates and their uncertainties for the four stream ecosystem metabolism models considered here while Table 6.6 compares the models' error metrics. The most complex model (saturating GPP and Arrhenius kinetics ER) performed similarly to the model with saturating GPP but constant ER on the calibration data (near equivalent RMSE, percent bias, and NSE), but the simpler model outperformed the more complex model on both the validation data and all data together (lower RMSE, % bias closer to zero, and greater NSE). Hence, we focus our analysis on the model with saturating GPP and constant ER (Figure 6.2).

For the complete days of monitoring (January 17th through 20th), cumulative daily ER was -4.97 g O_2 m^2 and for GPP varied from 4.86 – 5.04 g O_2 m^{-2}, reflecting the small variance in daily cumulative PAR for the period (67 to 70 mol photons m^{-2} d^{-1}). During the noon and dawn tracer injection, respectively, stream temperature averaged 21.6 and 17.5 °C while PAR averaged 2015 and 135 µmol m^{-2} s^{-1}. Consequently, the cumulative rates for the first 2.5 h of tracer passage were -0.52 g O_2 m^{-2} for ER but 1.18 and 0.13 g O_2 m^{-2} (noon and dawn injections, respectively) for GPP.

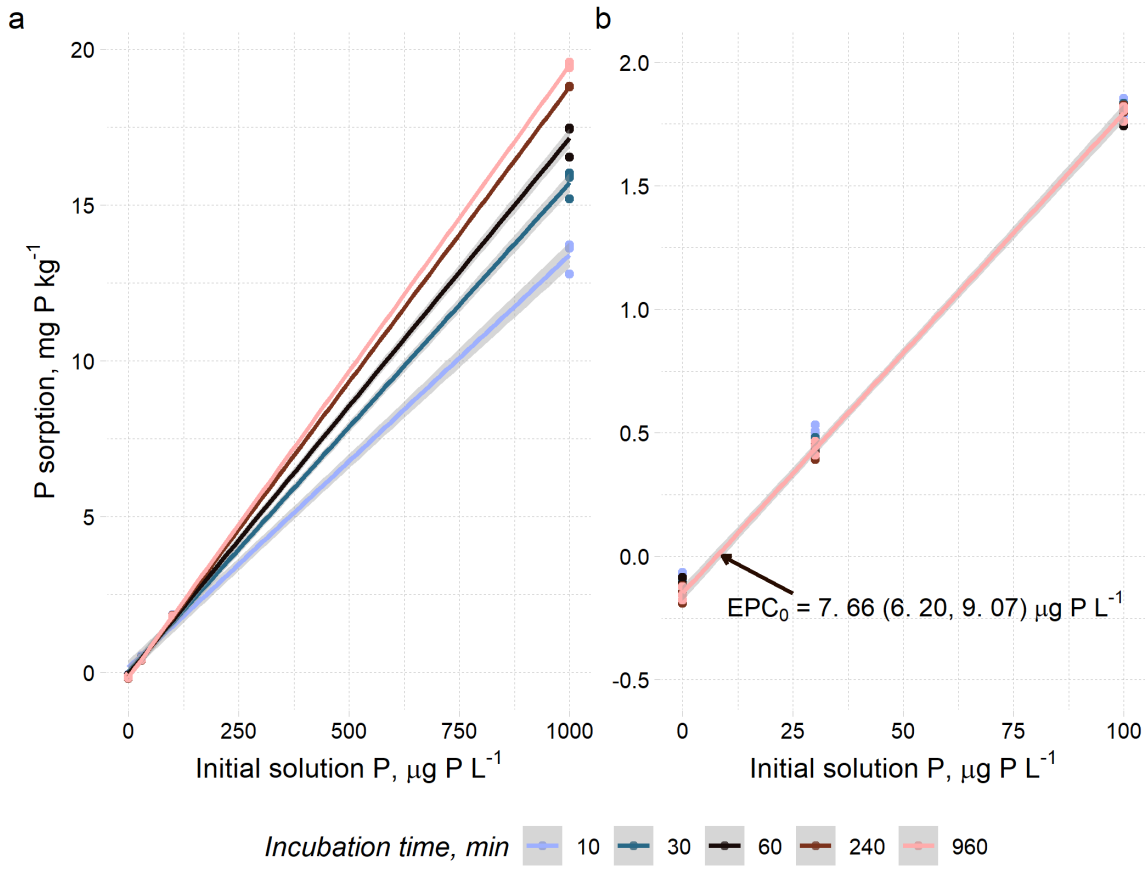

Figure 6.1 Sediment P sorption for benthic sediments. Longer incubation times allowed for greater sorption at greater concentrations (a) while, at much lower concentrations, sorption was similar and the EPC_0 (estimate only indicated for the 960 min incubation) was similar for times ≥60 min (b). Note the change in scales.

Table 6.5 Parameter estimates for the four stream ecosystem metabolism models considered. Parameter estimates are the MLE value from calibrating the model and the 95% confidence intervals are bootstrap estimates.

Model	**Parameter estimate [95% C.I.]**				
	P_{max}	α	E_b	ER_{ref}	K_{600}
	Units				
	g O_2 m^{-2} d^{-1}	g O_2 m^{-2} d^{-1} (µmol photon m^{-2} s^{-1})$^{-1}$	eV	g O_2 m^{-2} d^{-1}	d^{-1}
linear GPP, constant ER		5.8e-3 [5.4e-3 – 6.1e-3]		-4.6 [-5.0 – -4.4]	20.2 [19.2 – 21.5]
saturated GPP, constant ER	12.4 [11.8 – 13.2]	9.2e-3 [8.7e-3 – 9.7e-3]		-5.0 [-5.2 – -4.8]	20.4 [19.6 – 21.3]
linear GPP, Arrhenius ER		4.4e-3 [4.0e-3 – 5.1e-3]	0.47 [0.20 – 0.66]	-3.0 [-3.8 – -2.5]	13.6 [11.8 – 17.1]
saturated GPP, Arrhenius ER	10.4 [9.86 – 11.3]	7.3e-3 [6.8e-3 – 8.0e-3]	0.32 [0.21 – 0.42]	-3.6 [-4.1 – -3.3]	15.4 [14.2 – 17.0]

Table 6.6 Model performance summaries. For each partition of the dataset (calibration, validation, and all data; defined in methods), the model error metrics are root-mean-square error (RMSE) normalized to the observed standard deviation (SD), the percent bias (%), and the Nash-Sutcliffe efficiency (NSE).

Model	**Calibration (*n*=512)**			**Validation (*n*=386)**			**All data (*n*=898)**		
	RMSE:SD	Bias %	NSE	RMSE:SD	Bias %	NSE	RMSE:SD	Bias %	NSE
linear GPP, constant ER	0.207	-0.0839	0.957	0.607	0.624	0.630	0.381	0.198	0.855
saturated GPP, constant ER	0.160	-0.0686	0.975	0.477	0.286	0.772	0.298	0.073	0.911
linear GPP, Arrhenius ER	0.200	-0.103	0.960	0.637	3.37	0.593	0.395	1.28	0.844
saturated GPP, Arrhenius ER	0.158	-0.085	0.975	0.507	2.26	0.743	0.314	0.849	0.901

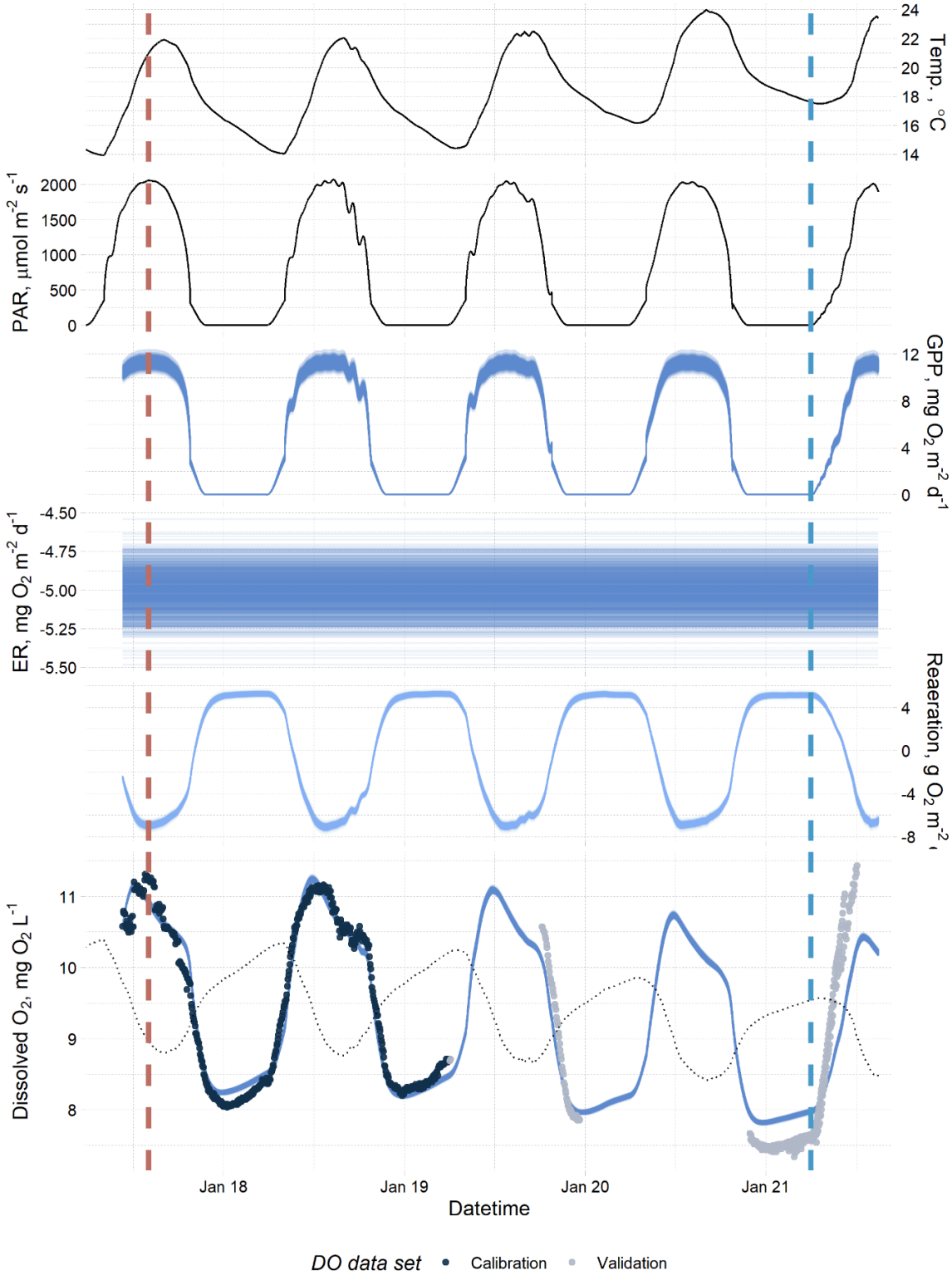

Figure 6.2 Stream ecosystem metabolism during the study period with times for the noon (red color, January 17th) and dawn (blue color, January 21st) injections noted with vertical dashed lines. The first two timeseries are for stream temperature and photosynthetic active radiation (PAR). The next three plots show the key components of the selected metabolism model on a streambed area basis: gross primary productivity (GPP), ecosystem respiration (ER), and reaeration (normalized to average depth). The bottom plot shows the measured dissolved oxygen (DO) as points, the saturation concentration as a dotted line, and the modeled DO in blue lines. The multiple lines in the bottom four plots represent a random subset of the bootstrap model fits (1000 shown); note that in this model ER was assumed invariant to stream temperature while GPP was a saturating function of PAR.

6.4.3 Phosphorus uptake during pulse injections

Injecting Br and DRP tracers concurrently at the same location but at two different times of day – dawn and noon – yielded slight differences in the transport and uptake of DRP (Figure 6.3). The peak concentrations for both Br and DRP declined longitudinally down the stream (moreso between B1 and B4 than between B4 and B5 despite similar distances) and receding limbs of the breakthrough curves had a characteristic exponential trend.

The $S_{w\text{-}add\text{-}dyn}$ for each observation on the receding limb declined linearly as $DRP_{tot\text{-}dyn}$ also decreased (Figure 6.4). The ambient uptake of DRP ($S_{w\text{-}amb}$) calculated for each monitoring station increased with further distance downstream (Table 6.7) but, at each monitoring station, $S_{w\text{-}amb}$ was less for the dawn injection (364 to 561 m) than for the noon injection (660 to 923 m). This indicates somewhat greater ambient areal DRP uptake (U_{amb}) during the dawn injection (0.39 to 0.61 µg P m^{-2} s^{-1}) than during the

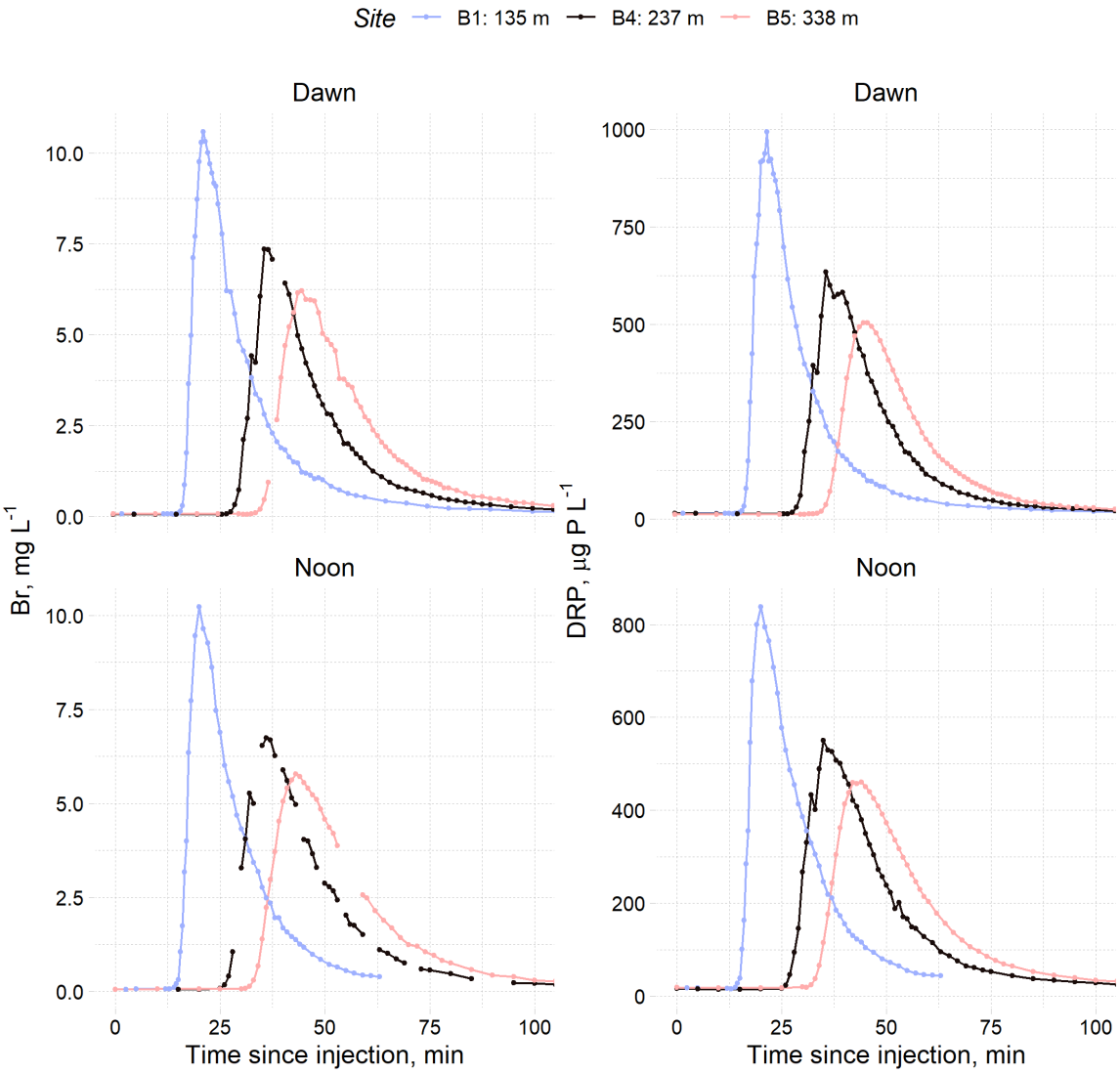

Figure 6.3 Breakthrough curves of bromide (Br) and DRP at the three monitoring stations during the dawn and noon pulse tracer injections. Breaks in the solid lines indicate missing observations due to analytical issues.

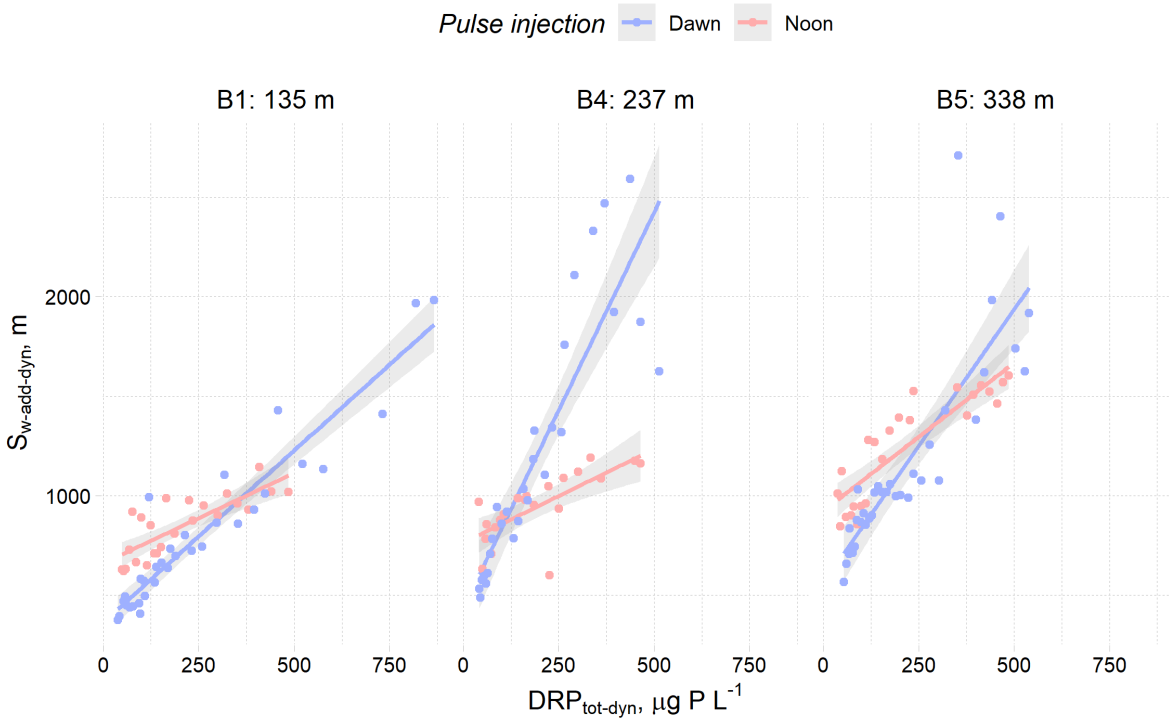

Figure 6.4 Dynamic solute uptake lengths ($S_{w\text{-}add\text{-}dyn}$) as a function of total dynamic DRP concentrations ($DRP_{tot\text{-}dyn}$) at each monitoring station for the dawn and noon tracer injections. Linear regressions for each group are plotted; the *y*-intercept corresponds to $S_{w\text{-}amb}$ for each dataset. Note that only data from the receding limb are analyzed here.

noon injection (0.37 to 0.41 µg P m^{-2} s^{-1}). However, while $U_{tot\text{-}dyn}$ increased with greater $DRP_{tot\text{-}dyn}$ for all locations and both injections, this rate of increase was greater for the noon injection than the dawn injection (Figure 6.5). A linear regression analysis of $U_{tot\text{-}dyn}$ (model summaries in Table 6.8) suggested lower intercepts (p<2e-16) but greater slopes (p<2e-16) for the noon injection compared to the dawn injection (Table 6.9). Slopes between sites for a given injection were similar.

Table 6.7 Ambient DRP uptake metrics determined via TASCC for both tracer injections.

Ambient uptake metric	**Units**	**Site**	**Injection**	
			Dawn	Noon
$S_{w\text{-}amb}$	m	B1	364	660
		B4	436	763
		B5	561	923
U_{amb}	µg P m^{-2} s^{-1}	B1	0.611	0.412
		B4	0.507	0.368
		B5	0.387	0.365
$V_{f\text{-}amb}$	mm min^{-1}	B1	2.56	1.50
		B4	2.29	1.38
		B5	1.89	1.21

6.5 Discussion

6.5.1 Phosphorus uptake under contrasting light conditions

In contrast to our hypothesis, ambient P uptake in the study stream did not increase with greater GPP (Table 6.7). In fact, the uptake metrics calculated through TASCC suggested that ambient P uptake was somewhat greater during the dawn injection than for the noon injection. This result supports sediment P sorption as a major contributor for P uptake in Kaituna River. We argue that this may apply to other streams where, previously, sediment P sorption has been overlooked.

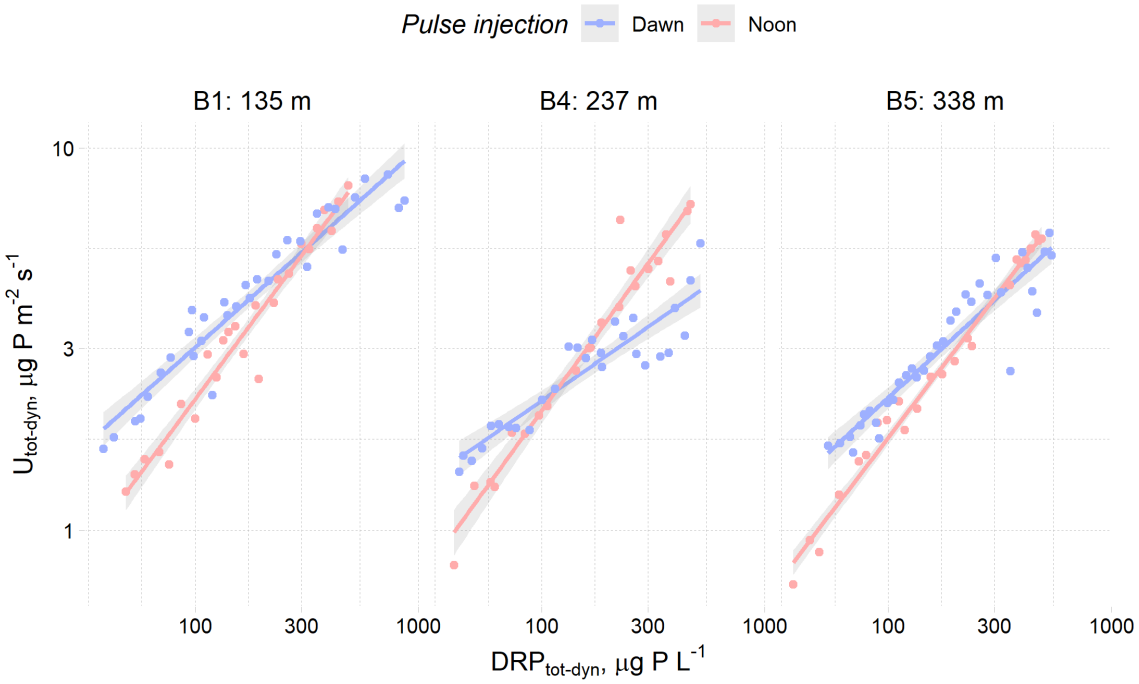

Figure 6.5 Total dynamic areal P uptake ($U_{tot\text{-}dyn}$) as a function of total dynamic DRP concentrations ($DRP_{tot\text{-}dyn}$) at each monitoring station for the dawn and noon tracer injections. Note that both *x* and *y* axes are on log scale. Linear regressions are shown for each group.

Table 6.8 Linear models for log-transformed $U_{tot\text{-}dyn}$. Model structures indicate the terms used in the model where ':' indicates interaction terms and '*' indicates full crossing (individual terms plus their interactions). DF is model degrees of freedom, RMSE is root mean square error, and AIC is Akaike Information Criterion. The ANOVA tests the model against the nested model (e.g., 2 against 1); the *p*-value is from the resulting *F*-test.

Model number	**Model structure**	**DF**	**RMSE [log(µg P m^{-2} s^{-1})]**	**AIC**	**ANOVA *F*-test**
1	log($DRP_{tot\text{-}dyn}$)	2	0.221	-24.6	
2	log($DRP_{tot\text{-}dyn}$) + log($DRP_{tot\text{-}dyn}$):Site + log($DRP_{tot\text{-}dyn}$):Injection	5	0.186	-76.3	p<2e-16
3	log($DRP_{tot\text{-}dyn}$)*Site + log($DRP_{tot\text{-}dyn}$)*Injection	10	0.130	-184.8	p<2e-16
4	log($DRP_{tot\text{-}dyn}$)*Site*Injection	12	0.127	-188.1	p=0.034

Table 6.9 Analysis of variance for model 3 (see Table 6.8) of $U_{tot\text{-}dyn}$. Interaction terms indicated by ':'.

Term	*F*-statistic	*p*-value
log($DRP_{tot\text{-}dyn}$)	2080	<2e-16
Injection	9.12	0.00295
Site	67.5	<2e-16
log($DRP_{tot\text{-}dyn}$):Injection	106	<2e-16
log($DRP_{tot\text{-}dyn}$):Site	1.65	0.195
Injection:Site	20.2	1.52e-8

The sediments studied here had potential for P sorption throughout the study, with an EPC_0 of about 8 µg P L^{-1} (Figure 6.1) compared to the ambient DRP concentrations of about 15 µg P L^{-1} (Table 6.2) and the peak breakthrough concentrations near 1000 µg P L^{-1}. This potential P adsorption may be realized given sufficient exchange of stream water between the water column and the hyporheic zone (Weigelhofer et al. 2018a; Simpson et al. 2021). The relevant physical and geochemical conditions for the sediment P buffer were largely constant between the dawn and noon injections: discharge (88 and 93 L s^{-1}; Table 6.1) was stable as were stream temperatures (17.5 and 21.6 °C) while stream pH was well buffered between ~7.4 to 8.5. Further, our laboratory P sorption measurements indicated a consistent linear increase in P sorption with increasing DRP concentration (up to 1000 µg P L^{-1}) and rapid kinetic rates, with the majority of sorption occurring within minutes (Figure 6.1). While we lack detailed measurements of the hydrological contact between the water column and fine benthic sediments, we expect that the coarse gravel surface layer, shallow stream depth, and brisk stream velocity were conducive to substantial hyporheic exchange during the study (Boano et al. 2014; Femeena et al. 2019). We suspect this hyporheic exchange was roughly similar between the two injections due to stable conditions and, therefore, the likelihood for sediment P uptake was also similar.

Unlike for sediment P sorption, the potential for periphyton P uptake likely differed between injections. Periphyton nutrient uptake often varies with PAR due to variation in GPP (Mulholland 1996; Quinn et al. 1997; Johnson and Tank 2009). During daylight, the periphyton in our study stream had considerable potential for P uptake. Mean periphyton biomass (50.6 mg Chl *a* m^{-2}; Table 6.4) was near the median of values compiled for ~300 streams by Dodds et al. (2002) and daily rates of GPP (near 5 mg O_2 m^{-2} d^{-1}; Figure 6.6) were in the upper quartile of rates compiled for 385 streams by Hall et al. (2016). Since sediment P sorption potential was similar between injections, this strong GPP provided the primary contrast in potential stream P uptake between our dawn and noon P injections. Indeed, instantaneous rates of GPP during the dawn and noon injections differed considerably: ~0 to 2 vs. ~10 to 12 mg O_2 m^{-2} d^{-1} (Figure 6.2). However, this increase in GPP did not correspond to any increase in ambient P uptake in the stream.

A key caveat here is that the periphyton were likely N-limited or N and P co-limited: DRP was relatively abundant but inorganic N species were below detection (Table 6.2). Streams with high autotrophic activity, like Kaituna River, tend towards N and P co-limitation (Elser et al. 2007; Welti et al. 2017). It is possible that the periphyton P uptake could have been more pronounced if background N concentrations were elevated (Piper et al. 2017), although some studies have shown a limited response in stream P uptake with N addition (Griffiths and Johnson 2018). Thicker, well-established periphyton

mats (as in the present study) may saturate their P demand at higher DRP concentrations (up to 50 µg P L^{-1}) due to diffusion limitations (Borchardt 1996).

Whether in experimental channels/chambers or in streams, some previous studies have hypothesized P uptake to depend on biotic uptake (i.e., periphyton) or on biomass metrics (Mulholland et al. 1997; Parker et al. 2018; Hanrahan et al. 2018), yet found limited evidence and missed the opportunity to emphasize sediment P sorption as a likely alternative mechanism. Our study as well as others propose instead that benthic sediment P sorption is major contributor for stream P uptake. For example, Martí et al. (2020) measured P uptake in a eutrophic, open-canopy stream across a gradient of light levels (PAR varying from 0 to ~1800 µmol m^{-2} s^{-1}). While NH_4-N uptake varied diurnally with GPP, the authors measured little variation in P uptake throughout the study and estimated miniscule contributions of primary producers towards P uptake (only ~2% of P uptake at the maximum GPP of 15 g O_2 m^{-2} d^{-1}) – they instead hypothesized the carbonate-rich benthic sediment to have had adequate P sorption potential to account for the uptake. Griffiths and Johnson (2018) measured P uptake via TASCC in a forested stream along with experimental manipulation of background NO_3-N concentrations. In their study stream, $U_{tot\text{-}dyn}$ increased with DRP concentration and varied between seasons (likely due to changing hydrological conditions) but this relationship as well as ambient P uptake metrics varied little between the NO_3-N treatments despite evidence for N and P co-limitation, suggesting P sorption as a dominant mechanism. Likewise, Wilcock et al. (2002) measured significantly greater P retention in an experimentally shaded stream reach compared to its unshaded reference reach: as this contrasts with the autotrophic P uptake hypothesis, they reasoned the increase in P retention of the shaded reach was most likely related to the greater transient storage – i.e. greater hydrological opportunity for sediment P sorption. These studies and our results clearly demonstrate that incorporating sediment P sorption alongside biotic P uptake pathways will improve our assessments of stream P cycling. Ignoring either mechanism can lead to incomplete reasoning about the cycling of P in streams.

Notably, while ambient P uptake did not vary with GPP as hypothesized, the relationships between $U_{tot\text{-}dyn}$ and DRP concentrations (as $DRP_{tot\text{-}dyn}$) suggested greater sensitivity in $U_{tot\text{-}dyn}$ to DRP during the noon injection compared to the dawn injection. Sediment P sorption rates would also increase across this range of concentrations (Figure 6.1), but it is conceivable that periphyton had some contribution to this greater sensitivity during peak production. As discussed above, N-limitation likely suppressed periphyton P demand, yet conditions in the stream during the noon injection were also highly conducive for N fixation by the periphyton mat (large periphyton biomass, low ambient N, warm temperatures, and high benthic irradiation; Marcarelli et al. 2008). Large periphyton biomass provides amble surface area and C substrate for colonization by other microbes, e.g., bacteria (Rier and Stevenson 2001; Battin et al. 2016), meaning greater C cycling during periods of high GPP may stimulate microbial activity (Rier et al. 2007) and so increase capacity for P uptake by the microbial communities within the periphyton mat (Matheson et al. 2012). Indeed, our study stream likely supported high overall microbial activity: the near 1:1 relationship between GPP and ER (Figure 6.6) suggest an intense cycling of autochthonous C

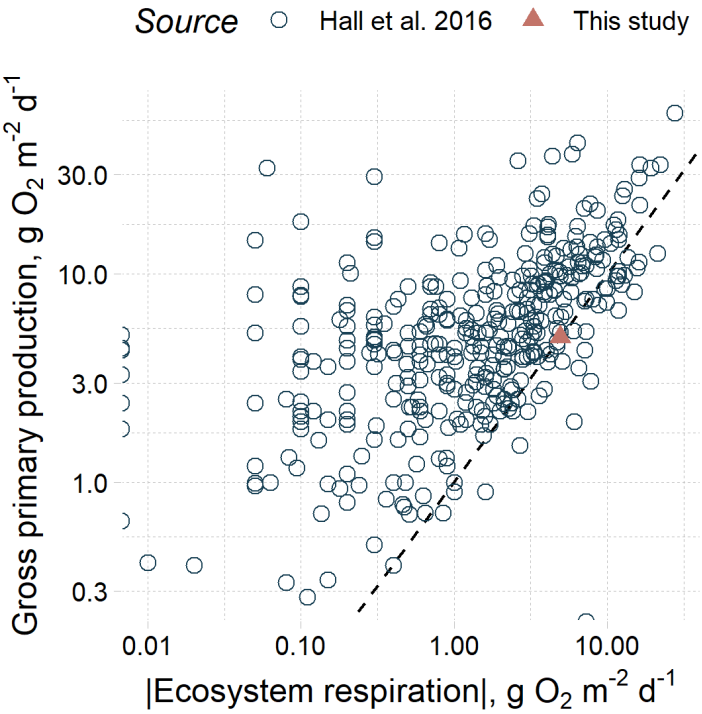

Figure 6.6 Comparison of daily whole-stream metabolism for the Kaituna river (this study) and for a variety of streams reviewed in Hall et al. (2016). Note that both axes are $\log_{10}$ scale. The dashed line is the 1:1 line.

(Finlay et al. 2011; Hall et al. 2016). Overall, while ambient P uptake did not vary, this greater microbial activity during high GPP may have provided a greater *capacity* for periphyton P uptake such as reflected in the increase in $U_{tot\text{-}dyn}$ vs. $DRP_{tot\text{-}dyn}$ slope during the noon injection (Figure 6.5). We are cautious in this interpretation, however, since there is considerable uncertainty in uptake metrics calculated through TASCC (Brooks et al. 2017); further replication is required to assess this trend.

6.5.2 Broader implications: sediment, periphyton, and in-stream P cycling

Pinpointing precisely the responsible mechanisms will improve our ability to predict P cycling (Haggard and Sharpley 2007; Dupas et al. 2019). Specific rates of P uptake can be rapid for either periphyton (Mulholland et al. 1994; Larned et al. 2004) or for sediment sorption (McDowell and Sharpley 2003b; Sposito 2004) but the opportunities for either to occur differ substantially. Storm events, grazers, and shading may prevent significant periphyton biomass accrual and thus P uptake in some streams (Biggs 2000; Dodds 2003) while restricted hyporheic exchange and saturation of sorption sites could limit sediment P uptake in others (Weigelhofer et al. 2018a; Simpson et al. 2021). These biogeochemical differences between sediments and periphyton – far from exhaustive – underpin our limited understanding of P cycling in streams. For example, both mechanisms may partly explain why, in comparison to either ammonium or nitrate, P cycling declines the most with increasing specific discharge (Ensign and Doyle 2006; Hall et al. 2013).

Benthic sediments are ubiquitous in streams but not always recognized for their important role in P retention (Hamilton 2012; McDowell 2015; Small et al. 2016; Kreiling et al. 2020). Simultaneously,

periphyton are an integral component of the stream ecosystem and, therefore, P retention (Dodds 2003; Welti et al. 2017). Knowing how streams attenuate P and what measurable drivers are important (e.g., light availability and sediment characteristics) can improve our predictions of whole-catchment P buffering. Disentangling within-stream P retention may then guide our expectations on management outcomes and sharpen our advice on mitigation actions (Meals et al. 2010; Jarvie et al. 2012).

6.6 Conclusions

Connecting in-stream P attenuation to P transport at the catchment scale is necessary to understand how our mitigation efforts perform. Here, we have separated two dominant, yet contrasting biogeochemical controls on baseflow P retention common to many catchments: periphyton P uptake and sediment P sorption. Despite considerable periphyton coverage and daytime GPP, differences in P uptake between injections performed at mid-day and early morning were minimal. Under the hypothesis that periphyton P uptake is a function of GPP, we argue that, alternatively, sediment P sorption may control most of the P retention in this stream. While this result is not generalizable – N addition could increase periphyton P uptake – this finding underscores the general importance of benthic sediment sorption to in-stream P cycling which, depending on the stream, may be a major form of P retention at baseflow.

Chapter 7
Synthesis and conclusions

7.1 Benthic stream sediments as a buffer for P

By examining multiple dominant forms for exchange of P (i.e., as phosphate and detectable as DRP) in streams, both biotic and abiotic, this thesis clarified how benthic sediments act as a *buffer* for P in streams.

This investigation was partly motivated by the strong correspondence between benthic sediment EPC_0 and in-stream DRP concentrations reviewed in Chapter 2. A comprehensive review and meta-analysis of the literature established how the relationship between these two variables – parameterized as a potential for the exchange of phosphate – is moderated by several environmental and experimental variables (e.g., sediment drying as studied in Chapter 3). Importantly, while sediments may have a potential for exchanging phosphate via sorption, several stream and sediment characteristics influence whether this potential is realized. For example, fine benthic sediments may prevent the hyporheic exchange necessary to transport phosphate in the water column to sorption sites within the benthic substrate. Hence, the apparent disparity between EPC_0 and DRP itself cannot tell the whole story as the sometimes counteracting causal mechanisms behind the phosphate exchange potential are still obscure. Chapter 2 identified that further research is needed to understand the sediment P buffer, e.g., how actual rates of sediment P fluxes vary under differing physicochemical conditions and how the sediment P buffer may interact with biota. Actual P exchange fluxes with the water-column, rather than just the potential indicated by EPC_0, will be critical for predicting how the sediment P buffer influences DRP concentrations at baseflow.

Recognizing the need to study the sediment P buffer in a wider geochemical context, Chapter 4 examined several abiotic mechanisms for P attenuation in streams at baseflow. Catchment geology was a critical control on P cycling in streams: the geology controls the water chemistry (e.g. alkalinity and cation activity) as well as the sediment chemistry (e.g., Fe oxide concentrations) pertinent to P attenuation. While (co-)precipitation of phosphate from the water column was an unlikely mechanism in these streams, all the sediments studied showed considerable P sorption capacity as well as large concentrations of labile P fractions (especially for the volcanic-basic geology). Sediment labile P concentrations correlated with in-stream DRP concentrations but only for streams with likely hyporheic exchange; lowland, spring-fed streams with large silt deposits (thus having minimal hyporheic exchange relative to stream discharge) had little correlation between sediment P variables and DRP. This observation reiterates an important facet of the sediment P buffer in streams: hydrological opportunity is required for the sediment P buffer to moderate DRP concentrations.

While the above work had a large focus on the abiotic component of P cycling, understanding biotic P cycling was also required to fully appreciate the sediment P buffer. Contrasting the sediment P buffer with biotic P cycling mechanisms proved to be an important facet of this work.

7.2 Disentangling biotic and abiotic controls on DRP

A central component of this thesis is that the sediment P buffer is driven by *abiotic* P sorption, which then is a major control of P uptake at the stream scale. The literature had previously either given P sorption little attention or even attributed what was likely the effects of P sorption to some other *biotic* process, e.g., microbial uptake. This thesis tested these hypotheses in Chapters 5 and 6.

Chapter 5 clarified some previous studies (see Table 2.1) on the microbial contribution to EPC_0 by also measuring the microbial biomass P. Sterilization procedures often increase EPC_0, but this is likely not due to a lack of microbial P uptake but rather due to the lysis of microbial cells whose biomass-P may then adsorb to the sediments. This ties in with the discussion given in section 2.5.3.1: exposing a sediment to more P, while holding its sorption affinity constant, will increase the EPC_0. Further, Chapter 5 tested the hypothesis that the microbial contribution towards sediment P uptake would increase when C and N are no longer limiting. Only N addition gave a marginal increase in EPC_0 (~5%), underscoring the considerable role of abiotic P sorption in the sediment P buffer.

With the addition of Chapter 6, this thesis isolated a likely sediment P sorption contribution to P uptake for one waterbody, the Kaituna River. The volcanic-basic geology of its catchment promotes high cation activity in the stream waters, yet the alkalinity is too low for Ca-related P cycling mechanisms such as calcite co-precipitation (Chapter 4). Sediment microbial biomass P is considerable in this stream (8 mg P kg^{-1} during the study of Chapter 5), likely supported by high autochthonous carbon inputs, yet the microbial component of P cycling was negligible compared to the (abiotic) sediment P buffer (Chapter 5). Finally, periphyton in this open-canopy stream represented a large potential pathway for P uptake: due to a likely N limitation, this component of P cycling in Kaituna River was also relatively small compared to the sediment P buffer (Chapter 6).

This investigation at Kaituna River was not exhaustive. Redox cycling of P-carrying Fe oxides is another major consideration for this stream; whether in the hyporheic zone (microbial reductive dissolution of Fe oxides) or in the water column (photo-reduction of Fe^{+3} during sunlight could encourage recycling of fresh Fe oxide precipitants), biogeochemical cycling of Fe is possibly a significant component of P cycling. Periphyton in the stream may increase their demand for P during periods of greater N availability (e.g., perhaps through a shift in species composition towards a community favoring N_2 fixation) – replication of this study in more P-limited ecosystems will provide a more severe test of the sediment P buffer hypothesis. In addition, more detailed measurements are needed to connect the biogeochemical reactivity of sediments with the hydrological opportunity of hyporheic exchange flows to fully elucidate the sediment P buffer in this stream.

However, this thesis demonstrates that understanding both biotic and abiotic mechanisms can refine our understanding of catchment P transport. Even if particular conditions vary (sediment composition, hydrologic features, etc.), benthic sediments carry an enormous potential for buffering P in many streams and should, at the very least, be considered alongside more well-studied mechanisms (e.g., periphyton P uptake). The sediment P buffer likely interacts with the biota in complex ways, such as by subsidizing microbial P demand, and so deserves more attention in stream ecology. Likewise, the strength and prominence of the sediment P buffer can determine how effectively it masks changing P inputs (e.g., as illustrated in Figure 2.8), and so deserves more attention in the research of catchment P transport.

This latter point is probably the most impactful consequence for P sustainability: the sediment P buffer is part of legacy P, or the persistence of past inputs of P in the catchment.

7.3 Future work on the sediment P buffer and stream P cycling

The measurement of EPC_0 itself is a topic in need of further research, as reviewed in Chapter 2. Chapter 3 highlighted the problems with drying sediments as a means for preservation prior to analyses. Drying sediments leads to biogeochemical changes which obscure the original P chemistry of the sediments: disturbance of organic matter, lysis of microbial biomass P (a considerable bias on its own; Chapter 5), and the ageing of metal oxides (i.e., transitioning from more amorphous to more crystalline oxides, which changes sorption affinity). However, other influences on EPC_0 measurement remain under-studied: solution pH and other background matrix chemistry, temperature, and appropriate equilibration time. Research on these topics could integrate with the tentative EPC_0 measurement approach developed in Appendix A.6, based on the review in Chapter 2, to improve the accuracy and generality of our measurements. It is unfortunate that a strict standardization of EPC_0 seems out of reach (Appendix A.6), but, more positively, a timely reminder of what to consider when measuring EPC_0 and related sediment P metrics should improve our future observations.

In addition to improving the handling and pre-treatment of sediments (Chapter 3), sediment sampling also needs to be improved or indeed the sediments studied *in situ*. Disturbance by removing sediments from the stream and through various measurements (e.g., prolonged periods of shaking) diminishes the relevance of such analyses to how sediments – as a matrix of mineral particles, organic matter, and biofilms – regulate P in streams. Isotopically labelled tracers may be an ideal approach, but stable isotopes for P other than the naturally abundant ^{31}P do not exist and radioisotopes (e.g., ^{32}P and ^{33}P) are usually too unsafe for use in streams. However, relatively new techniques such as measuring ^{18}O labelled phosphate (Jaisi and Blake 2014) may be a productive approach. Further, using a novel mixture of tracers which provide different types of information (e.g., the conservative Br tracer in Chapter 6 provides hydrodynamic information) may lead to new insights about reactive P transport in streams. For example, resazurin was proposed as a tracer that responds to metabolically active zones within the stream: resazurin, in the presence of aerobic respiration, reduces irreversibly to a biproduct, resorufin, both of

which are distinctly fluorescent (Haggerty et al. 2009). Injecting resazurin along with P and conservative tracers may help parameterize the contribution of both sediment sorption and microbial heterotrophs to stream P uptake and release.

In addition to the cycling of phosphate, the transport of sediments themselves is also critical for catchment-scale P transport. Sediments derive from variable erosion sources and are heavily attenuated throughout the stream network (Church 2002; Wohl 2015; Gran and Czuba 2017). Sediments act as a vector for P (e.g., Chapter 4) and therefore sediment transport regime is a critical research topic for the legacy P challenge.

Connecting the chemistry behind the sediment P buffer to actual sediment P exchange fluxes will require much more research, but such work would be crucial to quantifying rates of P cycling in streams. Phosphorus sorption to reactive surfaces is related to various surface complexation mechanisms (ligand exchange, surface precipitation), which are quantified by advanced surface chemistry techniques, e.g., Fourier transform infrared resonance (Arai and Sparks 2001; Parikh et al. 2014) and Mössbauer spectroscopy (Herndon et al. 2019). These are increasingly applied to soil environments to understand the mechanisms for surface P complexation, but similar research for stream sediments is lacking. Additionally, P *sorption* is better understood both in terms of chemical mechanisms as well as reaction rates compared to P *desorption*, but it is the latter which determines how well sediments maintain elevated DRP concentrations once upstream inputs decrease. Incubation experiments with specific tracers (see above), perhaps coupled with targeted spectroscopic techniques, could resolve questions around the nature of both sediment P sorption and other geochemical processes. Such questions include:

- Do sediments desorb P or is it more driven by, e.g., reductive dissolution of Fe oxides?
- How fast are the reactions for either P sorption or desorption in sediments at ambient conditions?
- Do changing redox environments, such as in the hyporheic zone, promote co-precipitation of phosphate with Fe oxide colloids which then carry P into the water column?

Relating to the idea of hydrological opportunity vs. biogeochemical reactivity in streams, comparatively little work has emphasized this concept in stream P cycling studies. For example, studies of hyporheic and groundwater flow paths coupled with dominant N reaction kinetics have enabled detailed accounts of in-stream N cycling (Zarnetske et al. 2011; Azizian et al. 2017; Harvey et al. 2018), but analogous work for P is lacking. Combining approaches from stream hydrology, stream ecology, and sediment (or soil) P chemistry should be explored as a way to build better causal models of P at baseflow. Likewise, as a final avenue for future research on the sediment P buffer in streams, both predictive and mechanistic models should be developed for stream P monitoring data. More high-frequency datasets (including of DRP) are becoming available (Bieroza and Heathwaite 2015; Vaughan et al. 2018) which can easily couple with high-frequency stream metabolism data (i.e., dissolved oxygen, temperature, pH, barometric

pressure, and light; Bernhardt et al. 2018). The latter can inform stream ecosystem metabolism models (e.g., for gross primary productivity and ecosystem respiration as in Chapter 6) which may then be used to test hypotheses (i.e., various model structures) on in-stream P cycling over daily and seasonal scales. Relatedly, a topic overlooked in this work is that of P transport through streams as various organic P species (Stream Solute Workshop 1990; Baldwin 2013) or even as phosphate complexed with dissolved organic matter (often via DOM-complexed cations, particularly Al and Fe; Hesterberg 2010). Indeed, the origins, cycling, and transport of DOM in streams is a burgeoning field (further aided by recent advances in high-frequency UV-vis and fluorometric data collection; Kaplan and Cory 2016; Vaughan et al. 2017; Wymore et al. 2018) and coupling this research with investigations on DOM-complexed P would be insightful for stream P cycling.

Formulating and testing these hypotheses against novel datasets can lead to better causal understanding as well as more accurate predictions of P availability and cycling in streams. Such predictions will prove critical in the future under a changing climate, where more frequent hydroclimatic extreme events – especially floods, droughts, and the timing of the two extremes – may exacerbate P losses to water bodies.

7.4 Closing remark: Sediment P buffer as legacy P

The element phosphorus was so named by its discoverer, the alchemist Hennig Brand, in reference to light: *phosphorus mirabilis* or 'miraculous bringer of light' (Jarvie et al. 2019). The connection to light is fitting. In our efforts to identify the sources of P in a catchment, which may eventually connect to the stream, we search for these 'light' sources as the astronomer does for distant stars or galaxies. The bane of anyone behind a telescope lens is *more light* – one will not be identifying any Messier objects during a full moon. In the case of P in the catchment, the light that obfuscates all other sources is legacy P.

Past inputs of P may be buffered by soils, vegetation, organic material, and – once in the stream – by sediments. This is called legacy P. For the stream, this legacy P may manifest as elevated EPC_0 (the 'level' of the sediment P buffer), which persists even if P inputs across the catchment are stopped. It is therefore important to be realistic about P management: determining the efficacy of mitigation or, on the other hand, the provenance of P measured in the stream will require (1) depletion of legacy P and (2) sufficient time for changes to propagate through to the stream. To target with certainty the likely P sources and their appropriate management practices, we must first account for the transient, legacy stores of P within the catchment including the sediment P buffer.

The sediment P buffer as legacy P is made even more important when one considers that our P monitoring – i.e., P concentrations in grab samples from the stream – is influenced by all the various buffers which P may encounter between the stream location and the initial point of P input in the catchment: the most *proximal* of these is likely the benthic sediment. In impacted catchments, the

sediment P buffer is the last ambient source of ‘light’ to our telescope, obscuring the luster of P sources upstream.

Appendix A
Supplementary Material, Methods, Results, and Discussion for Chapter 2

This appendix contains the following material related to Chapter 2:

- General procedure for determination of EPC_0
- Systematic search details
- Data collection details
- Statistical methods details
- Supporting Results
- How should we measure EPC_0? A recommended EPC_0 methodology baseline

A.1 General procedure for determination of EPC_0

Inheriting from soil science methodology (White and Beckett 1964; Taylor and Kunishi 1971 and references therein), the EPC_0 of a sediment is measured via a series of batch experiments. Some discrepancies behind the methodology are apparent in the literature (e.g., varying sediment to solution ratios, different background matrices, the concentrations used for the solutions, etc.). Here, we describe the method in general terms to illustrate how the data are generated:

Known masses of sediment (m_{sed}) are incubated with solutions of a constant ionic strength (i.e., the background matrix; deionized water, filtered stream water, $CaCl_2$, etc.) and known initial concentration of phosphate (c_i), typically ranging from 0 to >10 mg P L^{-1}, at a constant sediment to solution ratio (S, g g^{-1}; e.g., 1:20) and shaken, thus suspending and agitating the sediment, for periods on the order of (commonly) hours to (rarely) days; this time period only allows for a quasi-equilibrium as 'true' equilibrium (weeks to months) is not practical for these measurements (Barrow 1983a; Arai and Sparks 2007).

The solution is centrifuged and the filtered supernatant is analyzed for DRP (e.g., via the molybdenum blue method (Murphy and Riley, 1962)) as the final concentration (c_f), assuming all DRP in the solution is phosphate.

Net sorption is calculated as

$$q = \frac{(c_i - c_f)}{m_{sed}} V,$$

where V is the solution volume, or alternatively:

$$q = \frac{(c_i - c_f)}{S}.$$

After appropriate unit conversions, q may be expressed as ± mg P kg^{-1} sediment.

The measured sorption, q, is modelled on solution concentration (typically cf, although ci can be used as well; Haggard et al. 2004; Simpson et al. 2019) via an isotherm of the experimenter's choice (see Limousin et al. 2007). Linear isotherms are common, provided the concentrations used are close to the EPC_0; else, an L-curve type isotherm is typically used (Sposito 2004).

The point where the fitted model intersects the x-axis (q = 0 mg P kg^{-1}) is the EPC_0.

A.2 Systematic search details

Our systematic search followed best practice to ensure an unbiased dataset of paired DRP and EPC_0 observations (Côté et al. 2013) to study the phosphate exchange potential (PEP). We searched the Scopus and Web of Science databases to compile possible source papers. These two databases are favorable for natural sciences research and, when combined, provide excellent coverage of possible literature sources (Mongeon and Paul-Hus 2016). We focused somewhat more on the Scopus database as it allowed more fine-tuned searches (e.g., including abstract search results). In essence, we searched for papers that discussed sediment and P; included some mention of sorption or EPC_0; mentioned a lotic environment; and were in the English language. We did not restrict our search results in time. Both primary searches are from July 2018. We did not include the 'gray' literature (i.e., anything not published in a peer-reviewed journal).

Our search string for Web of Science, using the default search settings, was: TI=(sediment AND phosph*) AND TI=(stream OR river OR lotic OR riverine), with 346 results.

Our search string for Scopus was: TITLE-ABS (sediment W/5 phosph*) OR TITLE-ABS (phosph* W/5 (sorption OR adsorption OR desorption OR epc* OR equil*)) AND TITLE-ABS (stream OR river OR riverine OR lotic) AND (LIMIT-TO (LANGUAGE , "English ")), with 1713 results.

We found that our primary searches did not target some papers we had intended to include. These papers failed to appear in the initial Scopus search because the subscript in 'EPC_0' caused issues with the text search by concatenating 'EPC0' and any word that followed, e.g. 'EPC0value'. Therefore, we included

another Scopus search in September 2018 with the following string: ABS (epc0*), with 136 results. This auxiliary search resolved the issue and included the missing papers.

After removing duplicates, we screened the remaining 1807 results by title. Our reasons for exclusion based on title included obvious reasons (e.g., topic not related to sediment nor streamwater biogeochemistry) but also more specific reasons:

- Study focused on loads/fluxes and/or water-quality trends
- Simulation or modelling focus
- Study environment was floodplain, estuary, agricultural ditch, lentic wetland, reservoir, marine or other environment outside freshwater streams and rivers
- Microbial or ecological focus
- Policy or management focus
- Surface or solution chemistry focus (i.e., pure laboratory or theoretical chemistry studies)

After title-screening, we screened the remaining 433 results by abstract. Our reasons for exclusion here included the same reasons for exclusion by title above, but also:

- No indication of sediment P data
- Only discussion of a certain variable other than EPC_0, e.g., only P fractions

After abstract screening, we examined the remaining 119 papers at the full-text level. At this stage, the paper only needed to include concurrent benthic sediment EPC_0 and in-stream DRP data from a stream environment at baseflow conditions. Here, we relied on the authors' description to judge whether or not a stream was at baseflow as there was often little hydrological information available; generally, authors stated when a stream was not at baseflow. Including the aforementioned reasons for exclusion, we made two further reasons for exclusion. First, we excluded some papers where only an abstract was available (e.g., conference proceedings). Second, we excluded EPC_0 data for sediments size-fractionated beyond just <2 mm (e.g., <0.063 mm) rather than a bulk (<2 mm) sediment EPC_0 and excluded data from coarse sediment samples (e.g., gravel EPC_0) – we excluded these data for comparison purposes as (1) the majority of studies used a bulk fine sediment sample (<2 mm), and (2) coarse particles are more likely to contribute to apparent P uptake primarily through biotic mechanisms (rather than chemical sorption) not captured by EPC_0 (Lottig and Stanley, 2007). Additionally, only submerged, surficial sediments within the stream – those that would interact with the water-column at baseflow – were considered, e.g., floodplain or dry bank sediments (Kerr et al. 2010) were not examined.

Our search resulted in 48 primary papers for extracting data, and ultimately we used 45 of these (three papers requiring data from the authors were excluded as the corresponding authors were unreachable). Table A.1 gives the reference and basic information; further details are illustrated below.

The full search results and processing are recorded in the spreadsheet ('Search Master.xlsx') available at the Figshare repository reported in the main text.

Table A.1 The primary papers synthesized from the systematic review (45 total) with reference, country of study, and number of paired observations of DRP and EPC_0.

Reference	Country	Number of observations
Taylor and Kunishi, 1971	USA	11
Meyer, 1979	USA	2
Klotz, 1985	USA	4
Munn and Meyer, 1990	USA	2
Klotz, 1991	USA	26
Pailles and Moody, 1992	Australia	8
House and Denison, 1997	UK	16
House and Denison, 1998	UK	20
Haggard et al., 1999	USA	3
House and Warwick, 1999	UK	2
Bridgham et al., 2001	USA	2
McDowell et al., 2001	USA	4
House and Denison, 2002	UK	8
McDowell, 2003	NZ	5
McDowell et al., 2003	USA	23
Haggard et al., 2004	USA	16
Tamatamah, 2004	Tanzania	2
Jarvie et al., 2005	UK	82
Ekka et al., 2006	USA	80
Jarvie et al., 2006	UK	29
Popova et al., 2006	USA	10
Chaubey et al., 2007	USA	12
Haggard et al., 2007	USA	40
Lottig and Stanley, 2007	USA	3
Mosiej et al., 2007	Poland	6
Jarvie et al., 2008	UK	20
Stutter and Lumsdon, 2008	UK	12
Hoffman et al., 2009	USA	1
Lin et al., 2009	China	4
McDaniel et al., 2009	USA	128
Palmer-Felgate et al., 2009	UK	27
Lucci et al., 2010	NZ	9
Machesky et al., 2010	USA	12
Simmons, 2010	USA	12
Agudelo et al., 2011	USA	21
Zhang et al., 2012	China	17
Jalali and Peikam, 2013	Iran	17
Peryer-Fursdon et al., 2014	NZ	2
McDowell, 2015	NZ	76

Son et al., 2015	USA	10
Hongthanat et al., 2016	USA	4
Weigelhofer, 2017	Austria	8
McDowell et al., 2019	NZ	14
Roberts and Cooper, 2018	UK	36
Weigelhofer et al., 2018	Austria	96

A.3 Data collection details

DRP and EPC_0

For studies meeting our criteria, we extracted paired DRP and EPC_0 as our primary variables of interest. By design, this review excluded studies where either variable was missing. However, we made an exception for a few cases where DRP at the time of sediment collection was not reported but a representative average DRP value for baseflow at the site was available (for three studies, *n*=29).

Physicochemical variables

Along with pairs of DRP and EPC_0, we extracted the following physicochemical variables from each study: stream and sediment pH, total P, sediment fines concentration (defined as mass of particles <63 µm per total <2 mm sediment mass), bioavailable P via Fe strips (BAP; Sharpley 1993), exchangeable P fractions (via water or salt solution extraction, e.g., 1 M $MgCl_2$), the slopes of an isotherm equation (either from a linear isotherm, K_d, or from a Langmuir isotherm, k_{Lang}; data must be sorbed P plotted against final equilibrium concentration of P), the Bache-Williams index (BWI, see below; Bache and Williams, 1971), total C/organic C (via total elemental analysis), and organic matter content (via loss-on-ignition). We considered including P fractionation data since the chemical forms and bioavailability of P would provide relevant context to EPC_0, but ultimately dropped this variable due to inconsistent reporting and methodologies (Wang et al. 2013). We did attempt to include the 'exchangeable' P fractions (such as a water extraction step), if available.

Each numeric variable was converted to a consistent unit (e.g., mg P L^{-1} for both EPC_0 and DRP). This included BWI as well, since the choice of units and base of the logarithm influence the index value (Bache and Williams, 1971) – we converted (where possible) reported BWI values to units of (mg P kg^{-1})/$\log_e(C_{equil})$, where C_{equil} is the final solution P concentration in µg P L^{-1}.

Several studies reported either organic C content (OC) or loss on ignition (LOI, generally at 550 °C) as a proxy for organic matter content, but never both. The two variables are closely related (Pribyl 2010), and we applied a simple empirical conversion developed by Ball (1964):

$$OC = 0.458 \times \text{LOI} - 0.4$$

where OC is in % organic carbon and LOI is in % mass lost on ignition. The distributions of original OC data (*n*=167) and the OC derived from the above equation (*n*=145) were similar (mean ± SD; 0.78

± 1.46 and 2.39 ± 2.66 %, respectively). This synthetic OC variable was kept as a moderator variable in our analysis.

Site characteristics variables

We included catchment characteristics for each stream, where available. Catchment area was converted to ha basis. Land uses were included as given, with pastoral and arable uses combined as 'agricultural'; native and exotic forestry combined as 'forested'; cities, townships, and related terms combined as 'urban'; and all other uses classified as 'other unintensive', e.g., wetlands.

We also recorded any presence of a point-source upstream of the stream sites (primarily wastewater treatment plants). Many authors included such information, but where absent, we made no assumption on whether a point-source was upstream (left as missing).

Methodological variables

Since EPC_0 methodologies differed considerably across studies, and these differences likely contributed to between-study variance (Mengersen et al. 2013b; Nakagawa et al. 2017), we took the opportunity to appraise the effects of experimental conditions on the observed phosphate exchange potential (PEP). These methodological variables included: sediment pre-treatment (i.e., any type of drying), solution background used (see below), and equilibration time. Generally, authors stated explicitly what experimental conditions were used for their measurements. Therefore, we otherwise generally assumed that: sediments were analyzed fresh, at room temperature (~20 °C), P solutions were in deionized water (DI H_2O) with no other solution amendment, and microbial treatment (a biocide) was not used unless otherwise stated. Equilibration times were more difficult to judge, so we did not assume any value if not stated (i.e., left as 'missing').

Since solution background and ionic strength are important for P sorption (Barrow and Shaw 1979; Barrow 1983a), we recorded the nature of the solution in each study. Most studies used a dilute salt (e.g., 0.001 M $CaCl_2$), for which we calculated the ionic strength to make solutions comparable (although this ignores the effect of specific ions; Barrow and Shaw, 1979). Some studies used environmental waters (filtered stream or groundwater). Where available, we estimated ionic strength of these solutions by converting reported specific conductivity to ionic strength via Griffin and Jurinak's empirical relationship (Griffin and Jurinak 1973). Failing this, the solution ionic strength was unobserved.

Methodology for the moderator variables

It was not possible to account for every detail when collating the moderator variables. We suspect that there is variation for some measurements due to the specific method used. For example, sediment total P analyses are known to vary with digestion methods (Olsen and Sommers 1982), which were inconsistent across the studies in our review. However, we used the measurements as reported since these measurements correlate with each other and provide information (albeit imperfect) to the analysis.

Table A.2 Variables extracted from the systematic search results, the units or possible values, and the number of studies and percent total observations where the variable was available. Observations with differing units were converted to the unit indicated. Note that sediment refers to benthic sediments <2 mm.

Variable Name	Units or discrete values	Number of studies containing variable (% of observations)	Numeric variable summary (median [mean ± standard deviation])	Discrete variable summary (% of observations per level)
Primary Variables				
Stream water-column DRP	mg P L^{-1}	45 (100%)	0.042 (0.276 ± 0.766)	
Sediment EPC_0	mg P L^{-1}	45 (100%)	0.050 (0.202 ± 0.630)	
Stream and sediment physicochemistry variables				
Stream pH	standard units	21 (55.2%)	7.86 (7.85 ± 0.590)	
Sediment pH	standard units	10 (11.6%)	7.20 (7.10 ± 1.03)	
Organic C	g C kg^{-1}	11 (16.8%)	2.30 (7.44 ± 13.6)	
Organic matter (loss on ignition)	%	15 (39.4%)	2.20 (3.83 ± 4.55)	
Benthic sediment fines (<63 µm) concentration	g kg^{-1}	22 (54.8%)	141 (238 ± 262)	
Sediment total P	mg P kg^{-1}	18 (43.9%)	349 (593 ± 1040)	
Bioavailable P via Fe-strip	mg P kg^{-1}	3 (11.9%)	6.85 (24.3 ± 41.3)	
Exchangeable P	mg P kg^{-1}	12 (26.5%)	0.904 (2.41 ± 5.38)	
Bache-Williams P sorption index (Bache and Williams, 1971)	$\frac{\mathbf{mg\ P\ kg^{-1}}}{\mathbf{\log_e(\mu g\ P\ L^{-1})}}$	4 (1.2%)	2.45 (6.89 ± 6.10)	
Linear isotherm slope, K_d	$\frac{\mathbf{mg\ P\ kg^{-1}}}{\mathbf{(mg\ P\ L^{-1})}}$	12 (23.7%)	5.68 (242 ± 989)	
Langmuir isotherm affinity parameter, k	L mg^{-1}	4 (4.6%)	0.173 (0.412 ± 0.598)	
Site characteristics				
Watershed area	ha	24 (48.3%)	3840 (73700 ± 219000)	
Land uses	% of catchment	23 (51.8%)		
	Agricultural	22 (50.5%)	63 (56 ± 29)	
	Forested	20 (44.7%)	30 (41 ± 32)	
	Urban	15 (41.9%)	2 (4 ± 6)	
	Other unintensive (e.g., wetlands)	3 (6.2%)	24 (21 ± 8)	
Upstream point source influence		41 (76.3%)		
	No			67%
	Yes			33%
Methodological variables				
Sediment pre-treatment		45 (100%)		
	Fresh			92.4%

	Frozen or freeze-dried			1.4%
	Air-dried			6.3%
Solution correction		45 (100%)		
	DI H_2O			3.5%
	Dilute salt			54%
	Filtered stream or well water			42.5%
Solution ionic strength	mol L^{-1}	39 (71.5%)	0.006 (0.008 ± 0.007)	
Equilibration time	h	45 (100%)	18 (17 ± 14)	

Data overview and missingness

An overview of the dataset collected in this systematic review is illustrated in Figure A.1. Notably, most variables of interest were missing in at least one study. Joint observation of some variables was also poor (e.g., sediment pH and exchangeable P).

Several variables can be missing for a given PEP observation, which was generally a matter of whether or not the investigators measured or reported the variable. This form of missing data is termed 'missing at random' (MAR) in statistical literature (Rubin 1976; Schafer and Olsen 1998; Little and Rubin 2002) and is preferable to 'missing not at random' (MNAR; i.e., the probability of an observation being missing depends on the missing value itself), since, under MAR, imputation methods do not have to account for the missingness mechanism (Van Buuren 2018). An example of potentially MNAR data here would be left-censored DRP concentrations, i.e., values below detection, since the probability for being missing depends on the (unobserved) value itself.

We considered missing data in this review to be MAR. Generally, the data gathered here were not left-censored (one pair of DRP/EPC_0 in Ekka et al. (2006) were below detection and set to half the detection limit). We used data as reported in the literature but recognize that uncertainty likely increases as measurements approach detection limits. We do not attempt to propagate such uncertainty in our analysis due to complexity and sparse reporting of detection limits in the literature. In contrast, there were four EPC_0 values from one study that were right-censored (Roberts and Cooper 2018), which were from sediments downstream from point-sources and had EPC_0 greater than the highest batch solution P concentration used (1 mg P L^{-1}). As discussed, this constitutes MNAR data but we ignore this instance as it only relates to one stream from a single study in the whole dataset, which generally comprises MAR data.

Since we consider missing data in this review to be MAR, this allowed us to collect as much information as was feasible from the studies in the review and model PEP in a straightforward manner via multiple imputation.

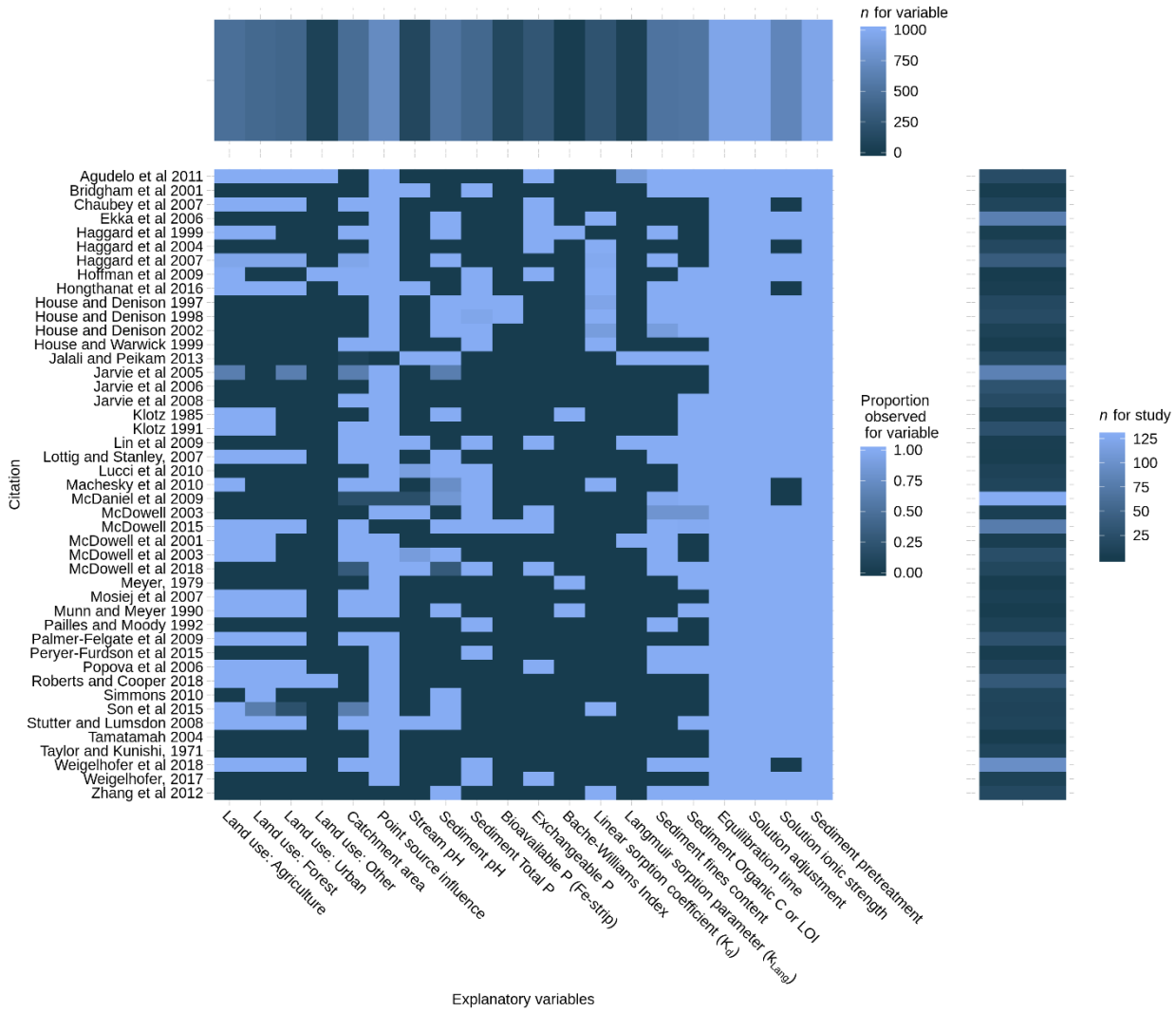

Figure A.1 Missingness of variables throughout the systematic review. Variables along the *x*-axis detail the proportion of data (pairs of DRP-EPC_0 observations) observed for a given study on the *y*-axis. The bars on the right and above of the missingness matrix give the number of observations for that study and for that variable, respectively.

A.4 Statistical methods details

Multiple Imputation: overview

Multiple imputation is a method for representing information in a dataset containing missing observations to make it amenable to complete-data analyses (Rubin 1976, 1996; Meng 1994; Van Buuren 2018). Multiple imputation fills in missing data in an informed manner (i.e., via predictive models based on related variables in the dataset and with consideration of the reason for missingness) and repeats this process iteratively and across separate chains to account for uncertainties inherent in the imputations. Following the generation of the multiply imputed datasets (numerous completed datasets containing the original observed data and unobserved data filled with one of the imputations), a statistical model can be estimated on each completed dataset and then pooled to a single, final model according to Rubin's rules (Little and Rubin, 2002; Van Buuren, 2018), thus providing parameter estimates that account for uncertainty not only in the data but due to imputation as well. The essential

steps in multiple imputation (Van Buuren and Groothuis-Oudshoorn 2011; Van Buuren 2018) are: 1) recognition of the missingness mechanism for each variable, 2) definition of predictive models to impute each missing variable, while preserving relations in the data, including which variables are used as predictors, 3) generation of multiply imputed datasets, and 4) pooling of statistical models across all imputations. Steps 1-3 are discussed here, while the 4th step (pooling) is discussed in the following section.

As discussed above, we considered the data to be MAR. That is, missing values of any one variable, given the rest of the data, come from the same distribution as the observed values (Rubin, 1976). The MAR assumption means the mechanism of missingness is ignorable for the multiple imputation process (Rubin, 1976; Van Buuren, 2018).

We used multiple imputation by chained equations (MICE) via the 'mice' package (Van Buuren and Groothuis-Oudshoorn, 2011). We defined an imputation model for each variable with missing data, utilizing all other information in the dataset, to generate *plausible* imputations to best *represent* the data. For example, if sediment organic carbon were missing, sediment fines content may provide a reasonable bound to our imputations, as greater organic matter tends to accumulate in sediments with finer texture (Findlay 1995; Tank et al. 2010); the inverse case is also informative due to the covariation. The MICE algorithm sets up *m* independent chains for the *m* imputations. Each chain iteratively cycles through each variable's imputation model, as some predictors in the imputation model also need to be imputed themselves, hence the chained equations need to converge before the random draw of the imputation model is representative (Van Buuren, 2018). While few imputations are required for good inferences, we used 500 imputations with 200 iterations to ensure stability of results when forming our final multiply imputed dataset.

Multiple imputation: imputation choices

For the imputation models (Table A.3), we used methods from the 'mice' package (details given by Van Buuren (2018)) and from the 'CALIBERrfimpute' package (method 'rfcont'; details in Shah et al. 2014). These methods were: predictive mean matching (pmm), Bayesian linear regression (norm), Bayesian logistic regression (logreg), and random forests (rfcont). All methods here ultimately produce a distribution to make a random draw from. Only point-source influence (as a Boolean variable) was predicted via logreg; the remaining continuous variables were imputed via either pmm, norm, or rfcont. Imputation models were selected for each variable considering the nature of that variable, predictors of interest, the observed data available, and judgements on missing data. A flexible machine-learning method like rfcont predicts well, handles interactions and non-linearities, and regresses towards the mean, thus avoiding imputations outside reasonable bounds (e.g., a sediment fines concentration above 1000 g kg^{-1}). Similarly, pmm performs well, preserves the original range of data (including extremes), and was favorable for the land use variables, as they were reasonably well-observed (covering 0 to

Table A.3 Imputation choices for the multiple imputation procedure. Note that sediment pre-treatment and equilibration time were fully observed and so did not require imputation. Refer to Table A.1 for full variable explanation. The organic C variable shown includes original and estimated organic C (synthesized based on organic matter via loss on ignition). Variables $\log_{10}$-transformed are indicated. Imputation models are referred to with their code name as used with 'mice': norm, pmm, logreg, and rfcont (see above text for descriptions). Variable code names are given for brevity and to match the R code given. Note that PEP, DRP, and EPC0 may be used as predictors as well.

Variable	Variable code name	Imputation model
Stream pH	pH_stream	norm
Sediment pH	pH_sed	rfcont
$\log_{10}$ Corrected organic C	log_OC_synth	rfcont
$\log_{10}$ Fines concentration	log_fines	rfcont
$\log_{10}$ Sediment total P	log_TP_sed	rfcont
$\log_{10}$ Bioavailable P	log_BAP	norm
$\log_{10}$ Exchangeable P	log_ExP	norm
Bache-Williams P sorption index	BWI	norm
$\log_{10} K_d$	log_K_d	norm
k_{Lang}	k_lang	norm
$\log_{10}$ Catchment area	log_catch_area	rfcont
Land-use: agriculture	LU_agri	pmm
Land-use: forest	LU_for	pmm
Land-use: urban	LU_urb	pmm
Land-use: other	LU_oth	pmm
Upstream point-source	PS_influenced	logreg
Solution ionic strength	I_str	norm

100%) but needed to meet the condition of (roughly) summing to 100%. For some variables, we used the norm method, as it efficiently incorporates joint variability between variables and can impute values slightly outside the observed data range which was necessary for some sparsely observed variables. In all cases, imputation models were inspected with diagnostic plots of imputed and original data (e.g., to check for unreasonable imputations or inconsistent distributions and correlations). Additionally, models were pruned if convergence issues appeared as a result of poor information contained in the predictor variables.

Predictor variables in the imputation models were chosen on the principal of including as many relevant variables as possible (Rubin, 1996; Van Buuren, 2018) to minimize the bias in imputations and to preserve the correlation between variables: this latter point also applies to the response variable itself (PEP), and so PEP appears as a predictor in many of the imputation models (White et al. 2011). For some imputed variables (e.g., PS_influenced), however, few relevant predictor variables existed. A simplified visual of the predictor matrix is given in Figure A.2 while Table A.3 relates the variables to their code name and MICE imputation model.

Multiple imputation: performance

Chain convergence in the MICE procedure is illustrated in Figures A.3 and A.4. Generally, chains were well-mixed and ceased to trend by <50 iterations, indicating that, by 200 iterations, the between-imputation variability was stable and gave unbiased random draws for these variables (Van Buuren and Groothuis-Oudshoorn, 2011). Notably, some variables showed the expected inter-relation through their

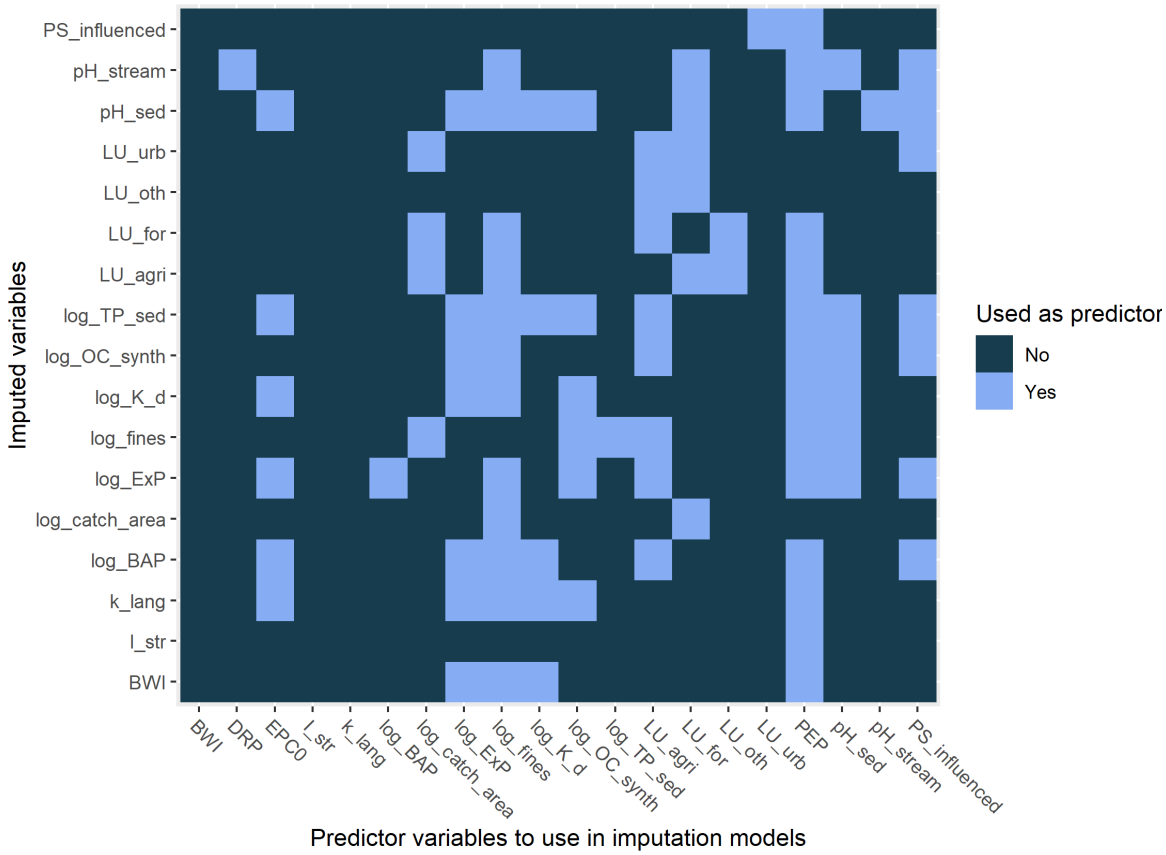

Figure A.2 Simplified predictor matrix used in the MICE algorithm. Variables on the *x*-axis indicate whether they are used as predictors in the model for the corresponding variable on the *y*-axis. For example, PS_influenced is predicted by a model including LU_urb and PEP while BWI is not used as a predictor in any imputation model. Variable names are as used in the corresponding R script (see also Table A.3).

chain convergences, e.g., how the LU_for and LU_agri chains converged in opposite patterns (initial trends) due to their interaction. Figure A.5 illustrates the imputation behavior with 100 of the imputed (completed) datasets. Importantly, Figure A.5 shows that the imputations reasonably preserve the covariation between variables, have stable distributions for well-observed variables, and preserve the original range in the data except when conservative extrapolation was necessary (via the norm imputation method).

Multiple imputation: pooling model fits for PEP across multiple imputations

Following multiple imputation, inference on the statistical model of PEP required combining each statistical model fit (m=500) into one pooled model via Rubin's rules (Rubin, 1987; Van Buuren, 2018). This step not only estimated a final model to base inferences on, but also estimated missing information metrics for the multiple imputations (Little and Rubin, 2002).

We pooled the GAMs (see main text) from the m = 500 imputations in an approach akin to that of Crane-Droesch et al. (2013). As mentioned in the main text, pooled statistics for the parametric terms included: parameter estimates, standard errors, p-values, $\bar{U}$, B, T_{var}, RIV, and FMI (for more details, see: Rubin, 1996; Schafer and Olsen, 1998; Little and Rubin, 2002; Van Buuren, 2018). Here, it was necessary to

pre-specify the knot locations for the cubic-regression splines when fitting each GAM (Crane-Droesch et al. 2013).

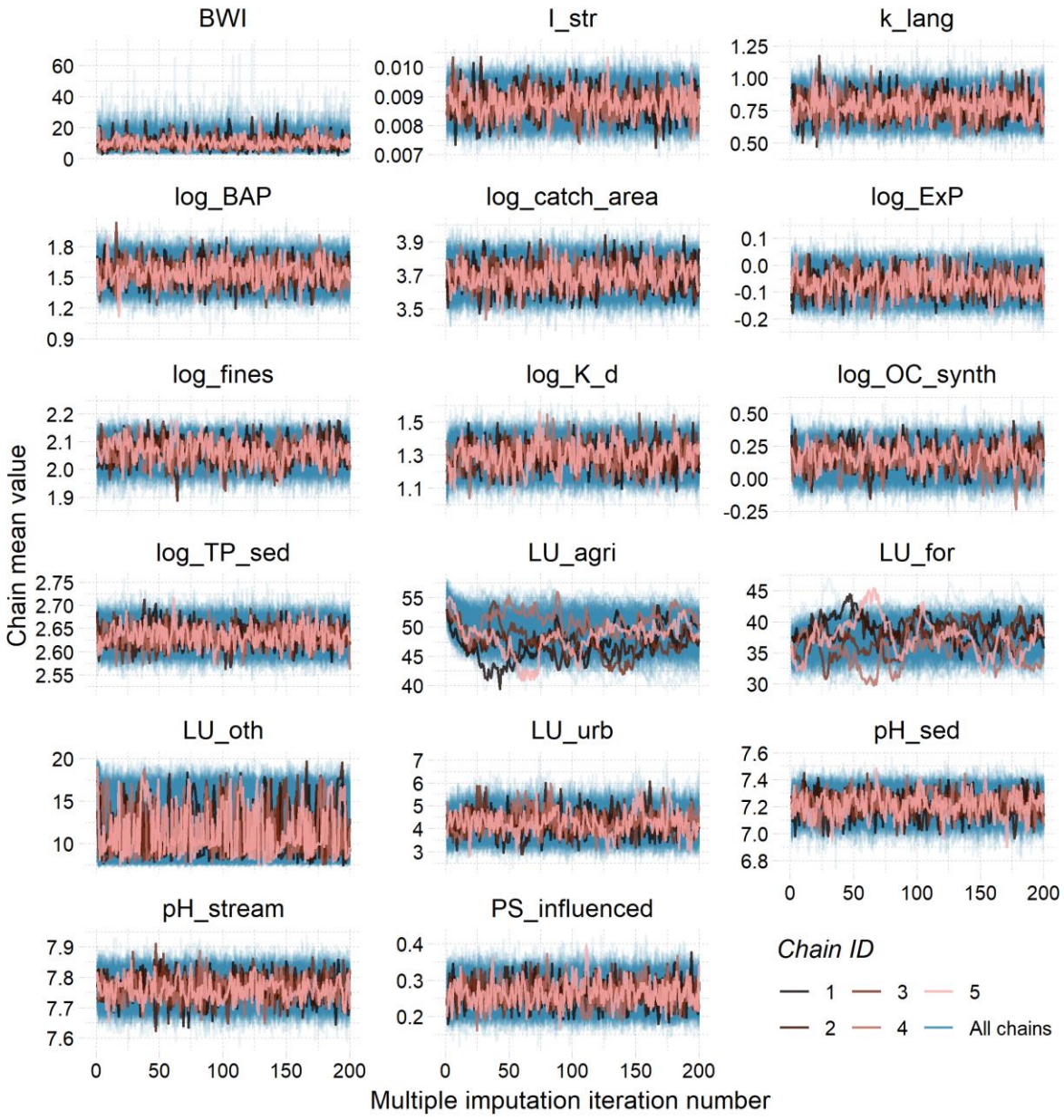

Figure A.3 MICE chain behavior with mean chain value (of only the imputed values at that iteration) plotted against iteration number. All *m*=500 chains are plotted in blue while the first 5 (random) chains are shown in the foreground as examples of individual chains. Variable names are as given in Table A.3.

While pooling of statistics for the parametric terms according to Rubin's rules is possible, the pooling for the smooth GAM terms, however, was not possible. For example, the *p*-values for these nonparametric terms cannot be derived from the pooled model (unlike for parametric terms) and the scalar *p*-values for the terms (see Wood 2013b) from each GAM cannot be pooled with Rubin's rules as *p*-values are not normally distributed (White et al., 2011; Van Buuren, 2018). However, we used an approximation developed by Licht (2010) and discussed by Van Buuren (2018) to pool *p*-values for the cubic regression splines (Wood 2013b) and for the random effects term (Wood 2013a). Note that the *p*-values for the smooth terms are themselves approximations (Wood 2013b), and Licht's procedure is an

approximation as well (Van Buuren, 2018) – accordingly, we interpret these *p*-values (denoted with *) with caution.

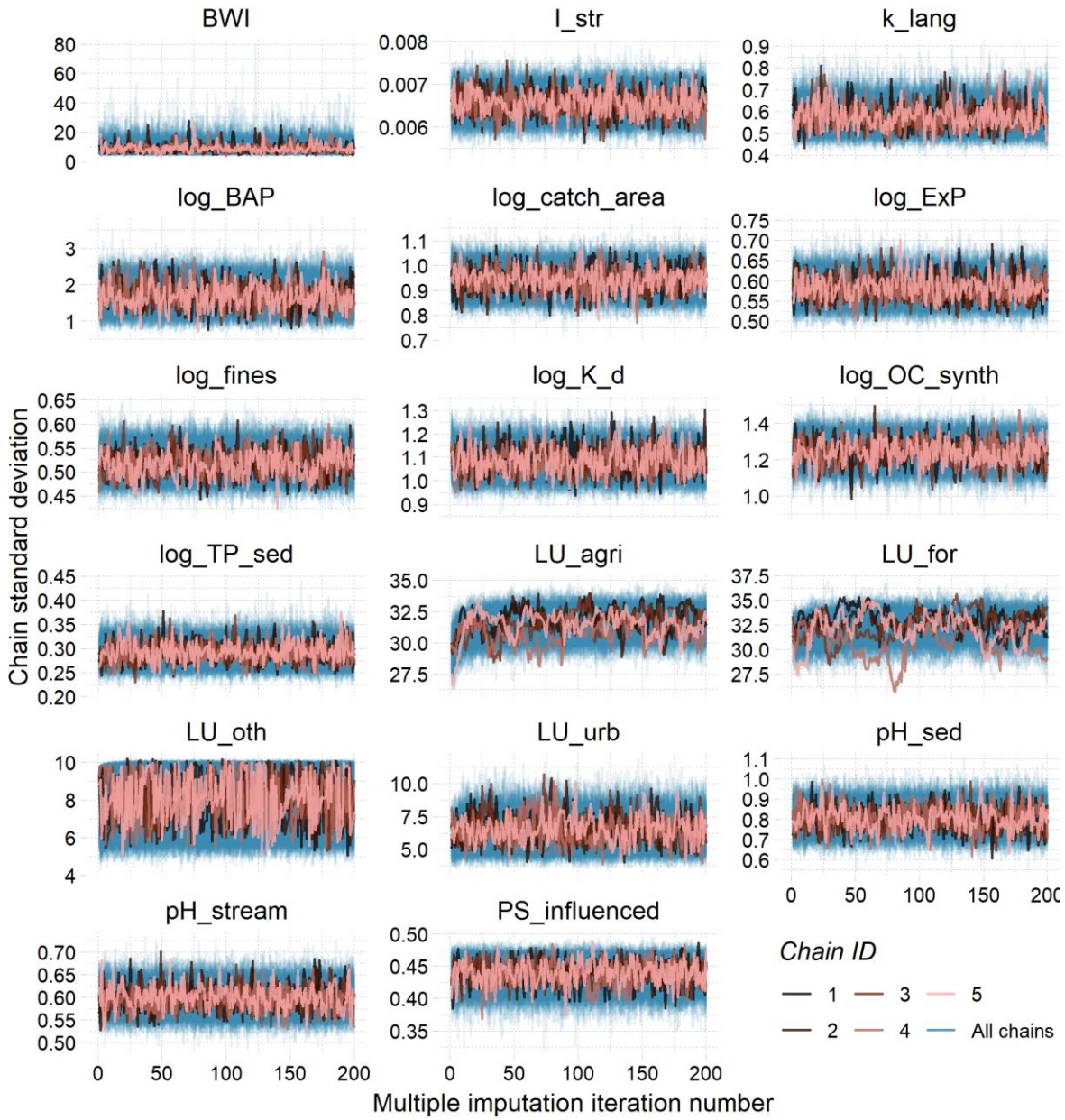

Figure A.4 MICE chain behavior with chain standard deviation (of only the imputed values at that iteration) plotted against iteration number. All *m*=500 chains are plotted in blue while the first 5 (random) chains are shown in the foreground as examples of individual chains. Variable names are as given in Table A.3.

Additionally, since direct estimation of within- and between-imputation variance for the nonparametric terms was not possible, we made an estimate of these missing information statistics (RIV* and FMI*) for the smooth terms as follows. First, we re-fitted all *m* GAMs with one of the smooth terms estimated instead as a linear term. We then pooled these models as above and derived approximate imputation statistics for that variable as normally for a parametric term. We note that the original parametric terms (parameter estimates and standard errors) were roughly comparable across the re-fitted GAMs, indicating that the linear approximation of the smooth term may have provided reasonable

approximations for B^* and $\overline{U}^*$; it is possible that these variances are inflated, however, due to missing the nonlinear characteristics in the data.

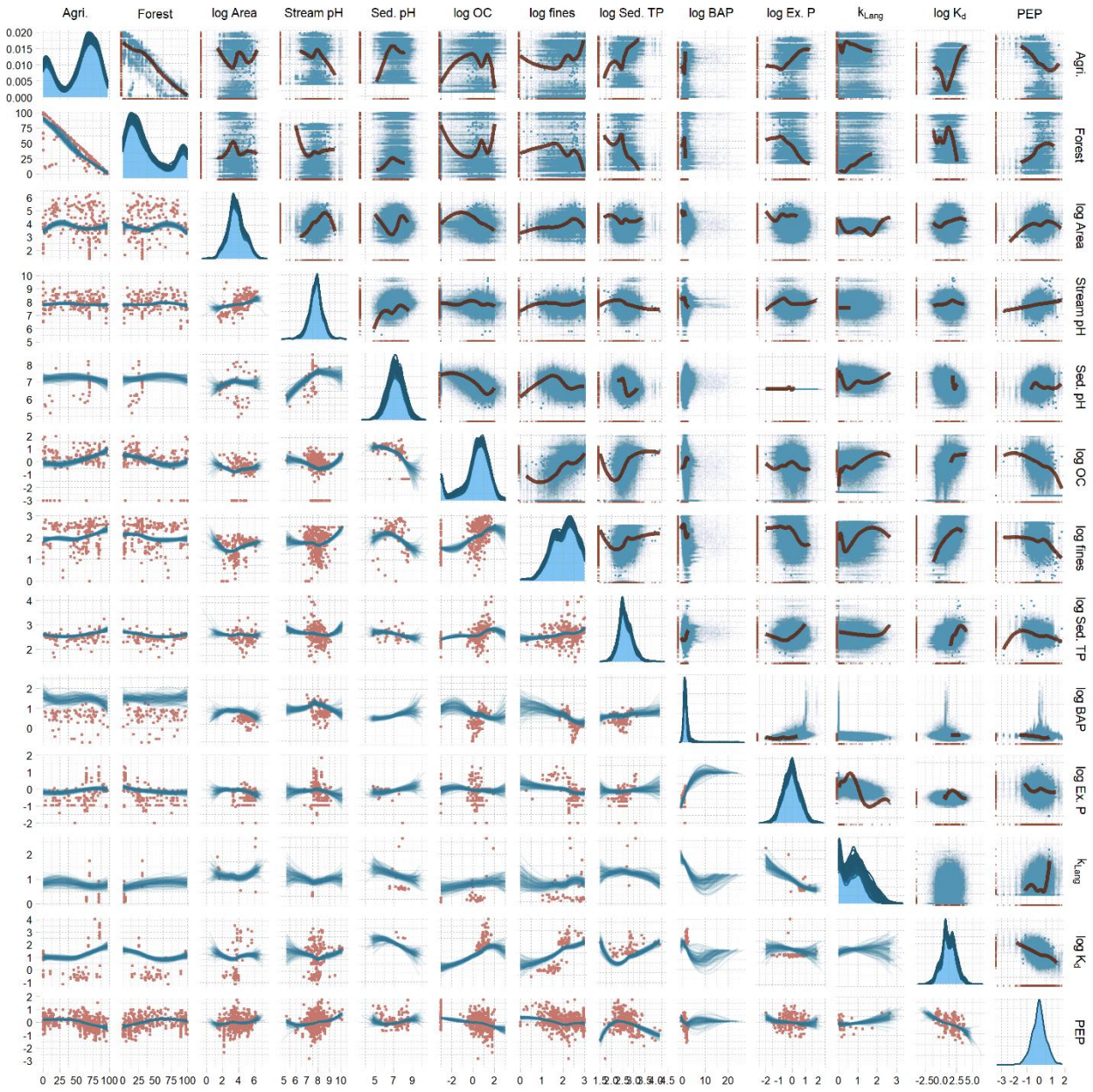

Figure A.5 Correlation matrix for the first 100 imputed datasets of our final multiple imputation approach for PEP and select variables. Across the diagram, red/pink coloring refers to the original observed data (with missing values) and blue refers to completed data (original observed data and with missing values filled in with imputations). The above half of the matrix shows the 100 completed datasets (out of a total of 500), with a single LOESS smooth to illustrate the shape of the original jointly observed data; rug plots along the margins show the marginal distribution of the observed data. The diagonal shows kernel densities for each completed dataset; note that PEP is fully observed and hence does not change meanwhile variables like log fines change mildly and poorly-observed variables like k_{Lang} fluctuate severely. The lower half shows the observed data (only jointly observed data can be plotted, hence plot data size varies); to summarize the relationship for the two variables in the completed datasets, one LOESS line is drawn per completed dataset (i.e., 100 lines are shown).

A.5 Supporting results

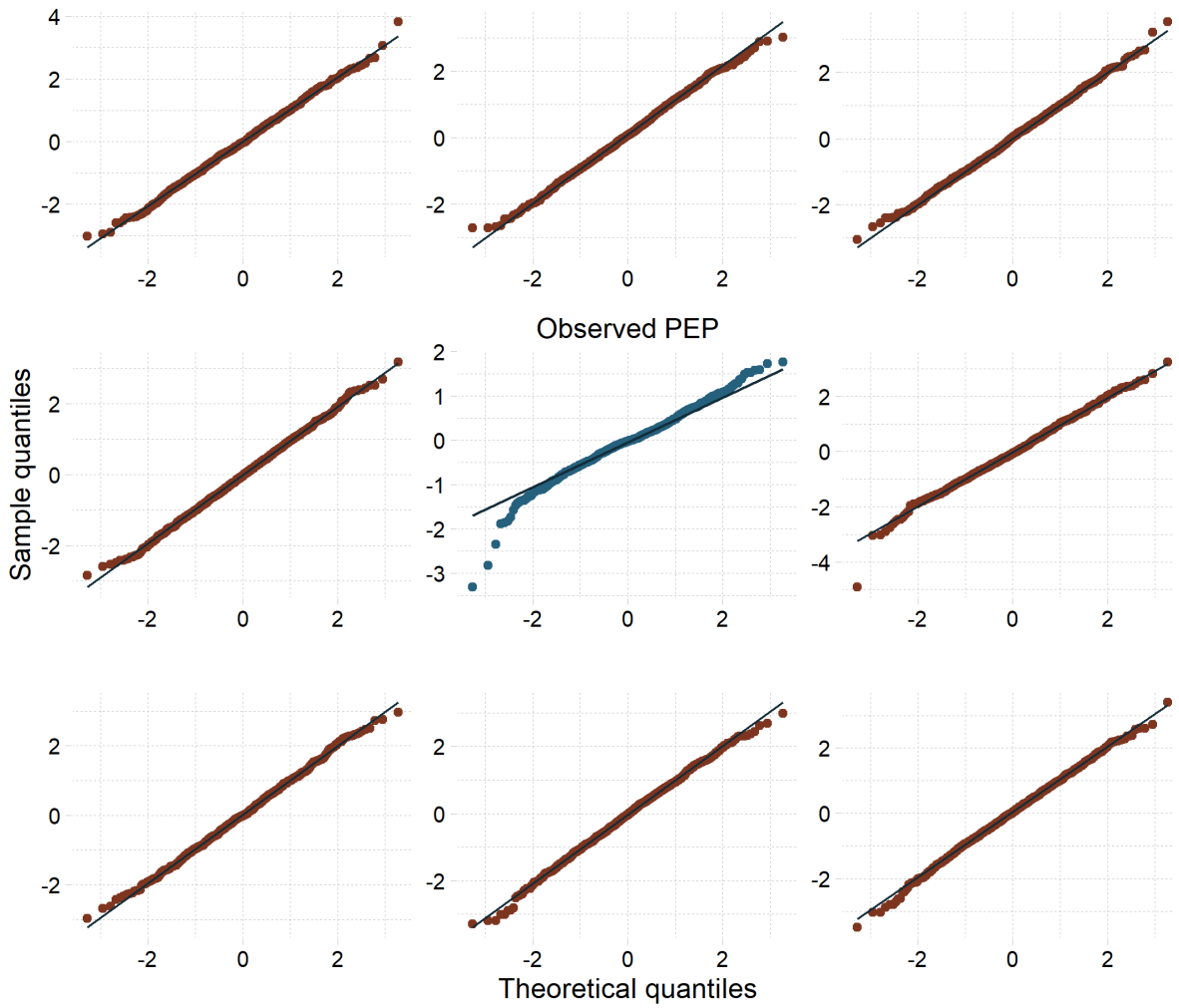

Figure A.6 Quantile-quantile plot of PEP (central plot) and 8 other simulated normal distributions of the same sample size (*n*=942) to illustrate quantile-quantile plots for normally distributed data. The quantile-quantile plot for PEP indicates a near-normal distribution but has slightly heavy tails in the distribution.

Table A.4 Summary statistics for the generalized additive models (GAMs) of phosphate exchange potential (PEP) fitted to the multiply-imputed datasets (m=500 imputations). All statistics for parametric terms are pooled estimates via Rubin's rules for pooling. Estimates (Est.) and standard errors (Std. Error) cannot be estimated for smooth terms (including random effects). Pooled p-values for smooth terms and random effects were approximated via the Licht-Rubin procedure. Multiple imputation statistics ($\overline{U}$, B, T_{var}, RIV, FMI) for smooth terms were approximated by re-fitting the GAMs with the given term re-fitted as a parametric term. $\overline{U}$ is within-imputation variance; B is between-imputation variance; T_{var} is total variance (accounting for between- and within-imputation variance); RIV is relative increase in variance due to nonresponse; FMI is fraction of missing information about the term. Bear in mind that all the estimates for the smooth terms are approximations only and should be interpreted with caution; these are denoted with *. 'Scale' refers to the remaining variance.

Term	**Est.**	**Std. Error**	***p*-value**	$\overline{\boldsymbol{U}}$	$\boldsymbol{B}$	$\mathbf{T_{var}}$	**RIV**	**FMI**
Parametric terms in GAMs								
Ionic strength (M)	-5.38	2.98	0.0712	6.31	2.56	8.88	0.406	0.291
$\log_{10}$ Equilibrium time (h)	-0.210	0.0437	2.08e-6	1.29e-3	6.25e-4	1.91e-3	0.487	0.330
Frozen or freeze-dried relative to fresh samples	0.362	0.157	0.0214	0.0200	4.68e-3	0.0247	0.235	0.192
Air-dried relative to fresh samples	0.241	0.0868	5.67e-3	5.96e-3	1.58e-3	7.54e-3	0.265	0.212
Point-source influence relative to no point source influence	-0.0986	0.0474	0.0380	1.35e-3	8.95e-4	2.25e-3	0.665	0.402

Smooth terms in GAMs*						
Term	*p*-value*	$\overline{\boldsymbol{U}}^*$	B^*	T_{var} *	RIV*	FMI*
Stream pH	2.29e-3	7.62e-4	8.25e-4	1.59e-3	1.08	0.523
$\log_{10}$ fines (mg kg^{-1})	0.0930	1.03e-3	1.89e-3	2.92e-3	1.83	0.649
$\log_{10}$ Exchangeable P (mg P kg^{-1})	0.0482	8.02e-4	3.53e-3	4.34e-3	4.41	0.818
$\log_{10}$ K_d (L kg^{-1})	6.08e-7	3.05e-4	9.64e-4	1.27e-3	3.16	0.763

Random effects in GAMs				
Term	Standard deviation	Lower 95% C.I.	Upper 95% C.I.	*p*-value*
Citation	0.00735	1.74e-3	0.00310	8.69e-7
Scale	0.464	0.443	0.485	

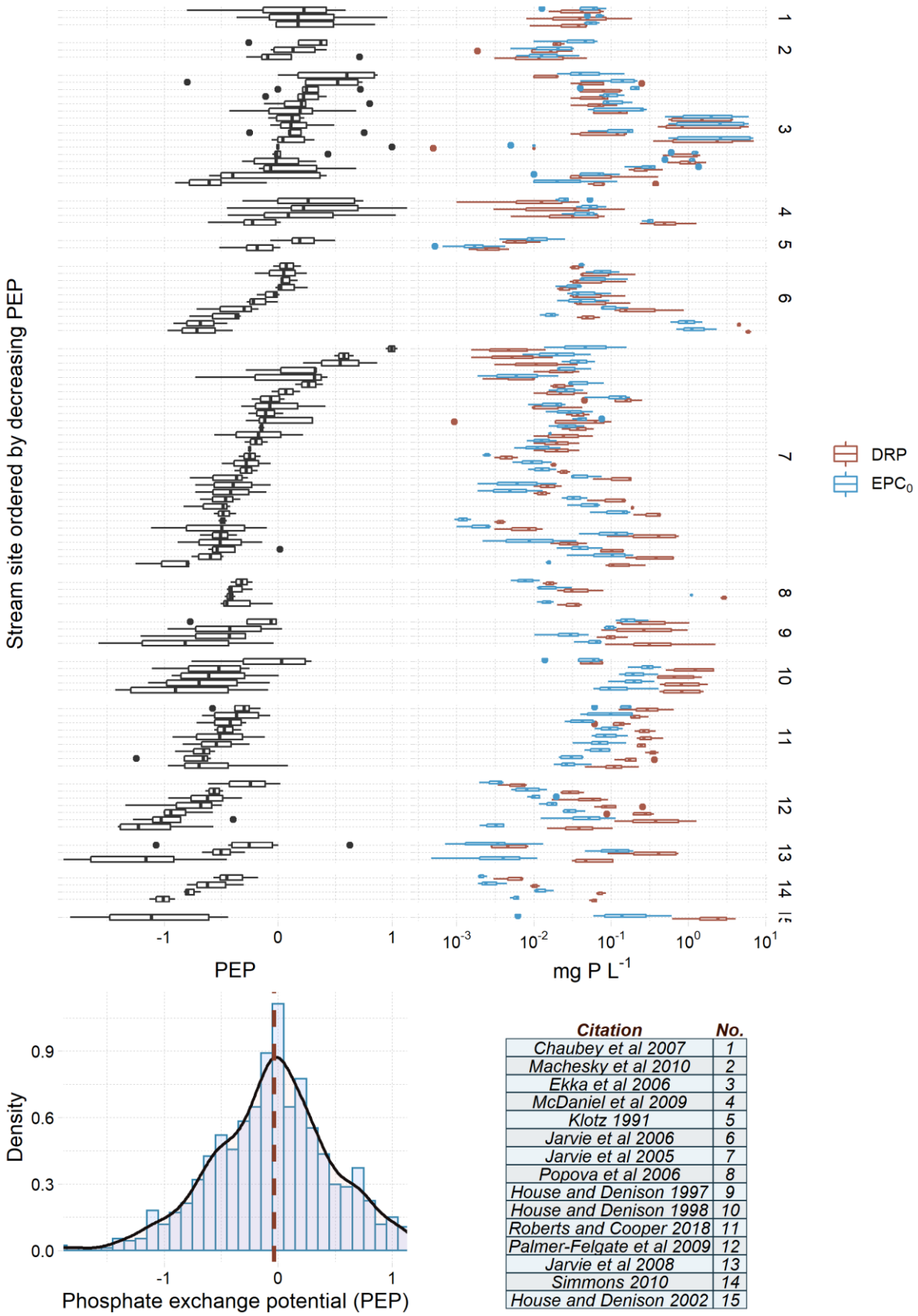

Figure A.7 Boxplots of PEP (left) and DRP and EPC_0 (right, log-scale) for streams with three or more observations. Data are arranged by decreasing site median PEP within a given study (see right side for reference number and bottom right for reference table). Below boxplots of site-specific PEP is the review-wide distribution of PEP for comparison (see also Figure 2.4 in main text); note that the full range is limited here for visibility.

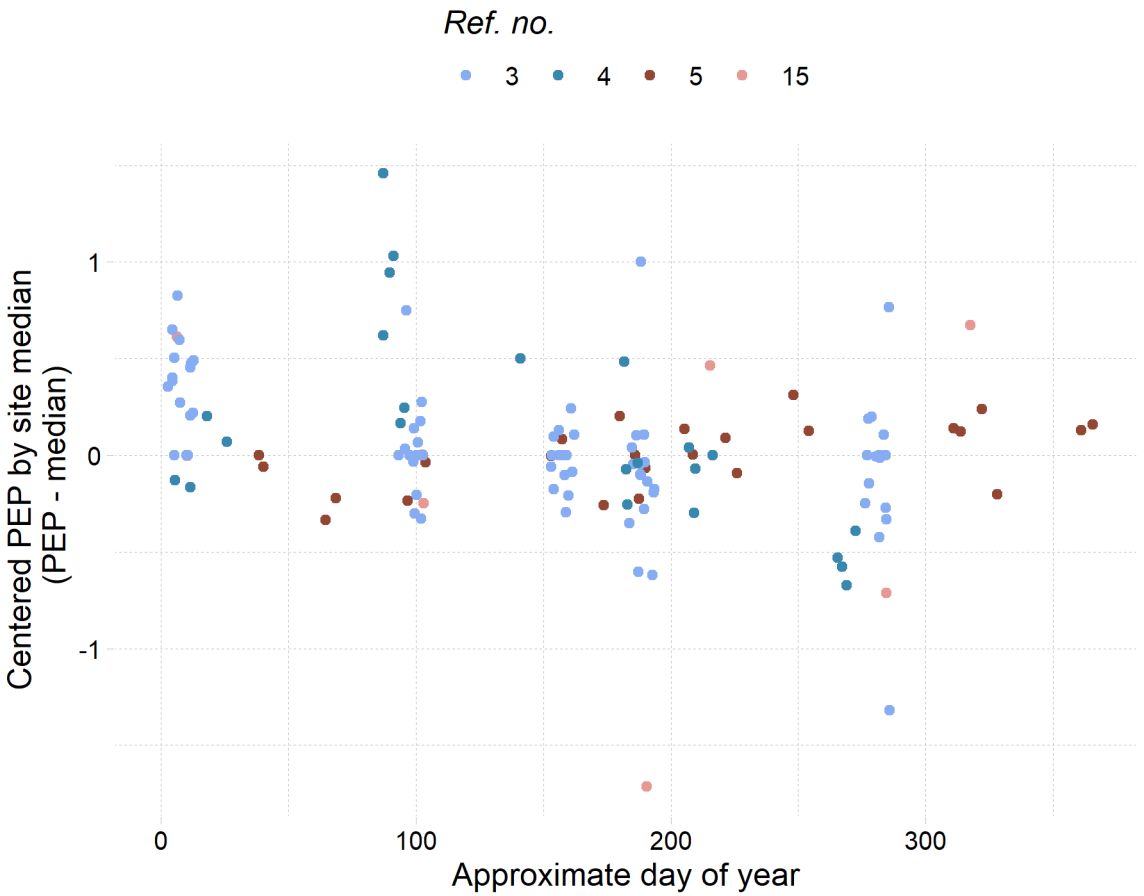

Figure A.8 Seasonality of PEP: four studies had stream sites with 5 or more observations over time; see the reference table in Figure A.7 for the reference number. Note that the PEP data here are centered by the median PEP for the given site for comparison purposes and that arbitrary jitter is added to the day of year for visual clarity. Note that references 3, 4, 5, and 15 shown here have seasonal data from 16, 4, 2, and 1 stream site(s) respectively.

A.6 How should we measure EPC_0? A recommended EPC_0 methodology baseline

> *"The equilibration solution should mimic the natural water in terms of ionic strength, specific cation, and pH if valid conclusions are to be made when comparing [EPC_0] and water column [DRP]."*
>
> *Klotz, 1988*

For EPC_0 measurement, we echo the consensus of Nair et al., (1984): there is no perfect method. This particularly applies to the choice of the background solution matrix. Like Klotz (1988), we desire the solution used to mimic the composition of *in situ* water. However, the partitioning of P between sediment surfaces and porewaters is dependent on a myriad of spatiotemporally dynamic factors (pH, temperature, etc.). This means that EPC_0 is best interpreted based upon the conditions of its measurement, which could resemble stream conditions but likely only over a limited amount of space/time. Hence, we cannot prescribe a standard method that would apply universally to streams. Future sediment chemistry work may enable us to extend point measurements of EPC_0 out to different conditions (e.g., estimate *in situ* EPC_0 based on a reference EPC_0 measurement and correction factors). For now, we urge careful control and monitoring of methodological variables followed by thorough reporting. Our recommended

approach below is guided by our objective for EPC_0 to be unambiguously measured, replicable, and reflective of *in situ* conditions. Careful modifications, however, may be made to suit specific research objectives.

Background solution matrix

We prefer preparing solutions in the laboratory but will also discuss the use of filtered environmental waters (below). Stock solutions can easily be made in the laboratory from D.I. water plus $CaCl_2$ and/or NaCl. We recommend first adjusting $CaCl_2$ concentrations to within the magnitude of Ca concentrations observed at baseflow (Lucci et al., 2010) then adjusting the final ionic strength of the solution to within an order of magnitude of that in the stream via NaCl. Average stream Ca activity is largely determined by the study catchment's hydrogeological setting and weathering status (Stumm and Morgan, 1996). As a useful approximation, Bluth and Kump (1994) broadly classified catchment denudation rates for catchments with lithologies dominated by (1) basalt, granite, and sandstones (low), (2) shales (medium), and (3) carbonate (high). Suitable Ca concentrations for streams draining such lithologies may therefore be: 0.5 mM $CaCl_2$ for low and medium denudation rates and 2 mM $CaCl_2$ for predominantly carbonate lithologies (i.e., karst). These values were derived from previous work in lithologies of (1) basalt, granite, and sandstone (Lottig and Stanley; Lucci et al., 2010), (2) shale (Klotz, 1988), and (3) carbonate (House and Denison 1998; Jarvie et al. 2005).

The higher concentrations suggested previously for EPC_0 (e.g., 10 mM $CaCl_2$; Taylor and Kunishi, 1971; Nair et al., 1984) were likely more relevant for soils (especially agricultural soils) rather than sediments. Unless the study stream is truly that calcareous, such high Ca concentrations should be avoided (Klotz, 1988; Lucci et al., 2010).

Following the addition of $CaCl_2$, the ionic strength of the solution should be adjusted to the same order of magnitude as that found in the stream at baseflow via addition of NaCl. Using the empirical relation from Griffin and Jurinak (1973), one can easily convert specific conductivity to ionic strength.

So, for example, consider an EPC_0 solution prepared for a stream draining a predominantly granite lithology whose average baseflow ionic strength (I) is around 2 mM. Starting with D.I. water, we may then adjust to 0.5 mM $CaCl_2$ (I = 1.5 mM) and 0.5 mM NaCl (I = 0.5 mM) to reasonably mimic Ca activity and ionic strength of the stream water.

Solution pH is also important but we lack information on how best to adjust pH. It is possible that sediment surfaces will buffer much of the pH changes during P sorption but addition of a pH buffer (e.g., Tris) near a representative, standard pH (e.g., pH 7 for many temperate, non-acid streams) could stabilize the EPC_0 measurement. It is unclear, though, whether such pH buffer chemicals interfere with P sorption, what pH buffer strength is necessary, and what pH should be targeted. Further research is necessary to settle how best to control pH for EPC_0. For now, unbuffered solutions may suffice.

As discussed in text, filtered environmental waters are problematic as background matrices for EPC_0. This background matrix may still be used for EPC_0 but only with due caution for: bias in ambient DRP due to non-phosphate yet 'reactive' P species; ambient DRP above the EPC_0; and changing physicochemistry of the water during storage (e.g., changes in pH and dissolved oxygen).

Regardless of the solution used, the ionic strength, pH, temperature, and Ca concentration at time of measurement should ideally be reported.

General incubation procedure

Keep sediments cool (4 °C) and in the dark prior to analysis within 1 to 2 weeks of sampling, preferably sooner. Using the appropriate background solution matrix (above), adjust DRP concentrations with a phosphate salt to three or more initial DRP concentrations. These concentrations should include a zero and values that closely bracket the expected EPC_0; if little prior information is available, concentrations above and below the in-stream DRP concentration by roughly one order of magnitude should suffice. Prepare incubations with fresh (wet) sediments at a fixed sediment to solution ratio (dry weight basis; 1:20 is suitable for many streams). Further, P sorption is subject to experimental error and so replicate incubations are encouraged. This enables not only the estimate of EPC_0, but its uncertainty as well (Simpson et al. 2019). Last, shake overnight (~16 h) at a stable temperature (e.g., room temperature of ~20 °C) to standardize the temperature effect; centrifuge and analyze for DRP, including occasional replicate analyses and quality-control checks.

Appendix B
Estimation of EPC_0 and its Uncertainty

B.1 Rationale for EPC_0 calculation and estimation of its confidence interval

Common calculations of EPC_0 involve either (1) simple linear regression (a linear sorption equation) of P sorption against equilibrium concentrations (i.e., solution concentration at the end of the incubation) and finding the *x*-intercept (Haggard et al. 2007; Hongthanat et al. 2016), or (2) including EPC_0 as a parameter to fit within an adsorption function for all batch equilibration points (House and Denison 2002; Jarvie et al. 2005). Both approaches are similar since the equations are essentially rewritten to include EPC_0 as a parameter to fit (see House and Denison 2000). For the case of the linear sorption equation:

$$q = K(c_{eq} - EPC_0)$$

where q is net sorption or desorption of P by the solid (mg P kg^{-1}), K is the affinity constant (L kg^{-1}), and c_{eq} is the equilibrium concentration of P (at end of incubation; mg L^{-1}). For a non-linear adsorption equation, e.g., the Freundlich isotherm:

$$q = K_f({c_{eq}}^n - {EPC_0}^n)$$

where n is a dimensionless constant ($0<n\leq1$) and the Freundlich affinity constant (K_f) now has slightly different units dependent upon n (mg^{1-n} L^n kg^{-1}).

These sorption equations imply that the desorption part of the function (when $q < 0$) is equivalent to the adsorption equation if it were reflected across the *x*-axis of a phosphate buffer diagram and reflected again across the line $c_{eq} = EPC_0$; this relation is not well-grounded given that desorption of P is not thermodynamically equivalent to its adsorption (Barrow 1983b, a; Miltenburg and Golterman 1998). Further, since the relation between q and c_{eq} is increasingly nonlinear further away from the EPC_0 (Barrow 1983a; Froelich 1988) it seems best to restrict observations to those closest to $q = 0$ to more accurately define the concentration where zero net sorption occurs.

However, sediments display the greatest P buffering capacity at solution concentrations near the EPC_0 (Froelich 1988). Consequently, sediment-P exchange experiments with c_i slightly above and below the EPC_0 will produce data containing both positive (adsorption) and negative (desorption) values of P sorption for a c_{eq} practically indistinguishable from the EPC_0 (notwithstanding experimental error and artifacts due to sediment heterogeneity). In other words, for the same *x*-coordinate, two different values of *y* are apparent. This issue arises because c_{eq} and q are not truly independent variables: sorption or desorption observations are derived from the observed solution concentrations (Barrow 2008).

This vertical gradient is difficult (or impossible) to explain with any continuous function as the slope is undefined. An example of this failure with experimental data (sediment T10 from Chapter 3) is given in **Figure B.1** . In some (but not all) cases where the sediment displays a strong buffering capacity, using c_{eq} in a linear regression yields a poor estimate (0.019 mg P L^{-1}). Even with removal of one extreme desorption point (from the 0 mg P L^{-1} treatment) with a final c_{eq} of ~0.025 mg P L^{-1}, the linear regression on c_{eq} results in an incorrect fit and EPC_0 (fit not shown; 0.016 mg P L^{-1}).

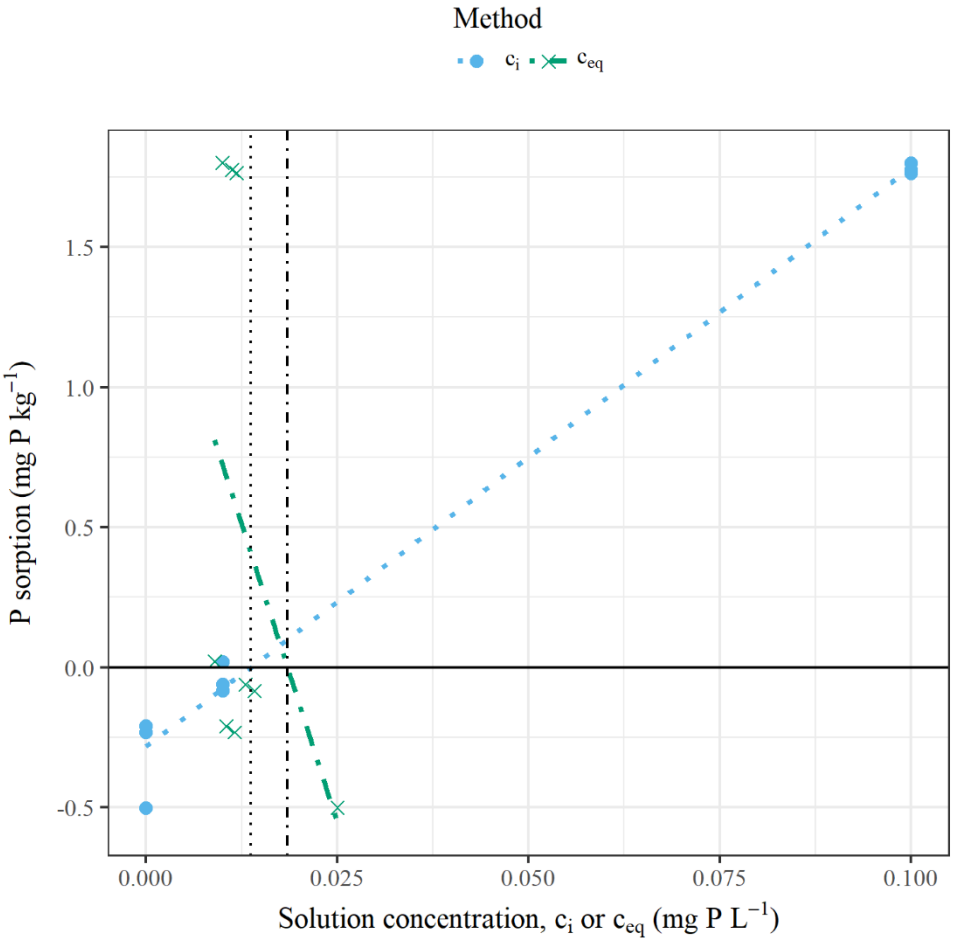

Figure B.1 Comparison of using initial solution concentration (c_i) and equilibrium concentration (c_{eq}) when calculating the point of zero net phosphate sorption (EPC_0) via linear regression for one sediment (T10; see Chapter 3); the alternative c_{eq} is the same data as c_{eq} but without the point at 0.025 mg P L^{-1}; vertical lines indicate the resulting EPC_0 value for a given method.

Throughout this thesis, sediment EPC_0 is calculated as the *x*-intercept of the linear equation of net P sorption on c_i instead of c_{eq} (e.g., Haggard et al. 2004). We consider this approach more appropriate as c_i is an independent experimental variable where, for one value of c_i, one response value (q) can be reasonably expected. However, we do acknowledge that q is still a variable derived in part from c_i and that the linear equation will be inadequate for sorption values at higher c_i (i.e., further from the EPC_0). Therefore, we only use observations of net desorption and the first triplicate set of points for a given c_i exhibiting only net adsorption (in this study, mostly the c_i = 0.1 mg P L^{-1} treatment). With this data, the linear sorption equation, $q = K(c_i - EPC_0)$, is fitted with nonlinear regression which allows for a direct

parameterization of EPC_0 and its uncertainty (i.e., 95% confidence interval) via a profile-likelihood method (see example 4 in Harding 1986). The R code and data used for these calculations is provided online at https://doi.org/10.6084/m9.figshare.6157772.v1.

We illustrate our simple approach on the same dataset as before in Figure B.1. In contrast to the estimate via c_{eq} of 0.019 mg P L^{-1} (0.016 mg P L^{-1} if the low desorption point is ignored), we now estimate EPC_0 as 0.014 mg L^{-1} with a 95% confidence interval from 0.009 to 0.018 mg P L^{-1}. Further, we suggest that future studies incorporate more experimental sorption points near the EPC_0 (if prior information on the sample exists) to determine EPC_0 more accurately and other sorption parameters (e.g., sorption affinity).

Appendix C

Supplementary Results for Chapter 3

C.1 Supplementary results for Chapter 3

Table C.1 Physicochemical characteristics of the sediments analysed in Chapter 3, organized by catchment. ASC is anion storage capacity; pH is for fresh sediments only; and 'NA' stands for 'not available'.

ID	Ca (g kg^{-1})	Al (g kg^{-1})	Fe (g kg^{-1})	Mg (mg kg^{-1})	Mn (mg kg^{-1})	K (mg kg^{-1})	Na (mg kg^{-1})	Zn (mg kg^{-1})	Total C (g kg^{-1})	ASC (%)	pH
Tukituki Catchment											
T1	4.50	10.8	19.8	4600	360	870	113	55	1.36	5.2	6.52
T2	5.00	10.6	20.0	4700	380	810	122	53	1.5	6.6	7.43
T3	4.50	10.2	19.6	4600	350	690	103	47	1.53	4.9	7.16
T4	4.40	9.2	18.6	4400	340	710	101	51	1.64	6.6	6.35
T5	4.50	9.8	18.6	4400	360	730	118	49	1.01	7.0	7.26
T6	3.40	8.1	16.5	3900	300	590	107	45	1.31	6.4	7.4
T7	4.42	14.0	17.8	4318	281	2498	220	52	1.22	5.3	6.89
T8	3.30	8.0	16.1	3800	300	640	122	43	1.52	4.1	7.29
T9	6.50	22.1	18.9	4676	315	3650	336	62	5.73	12.0	7.55
T10	3.10	7.3	14.6	3600	270	650	123	41	0.83	7.5	7.32
T11	3.30	7.7	14.9	3500	270	680	118	41	2.46	8.0	7.57
T12	5.40	8.6	16.0	3700	270	950	176	49	0.84	6.7	7.71
T13	5.96	20.9	17.9	4404	292	3822	347	55	5.06	11.7	7.76
T14	15.60	6.6	14.0	3200	260	780	124	39	4.93	9.7	8.09
T15	7.40	8.9	16.6	3700	290	940	178	51	2.41	9.5	7.87
T16	6.42	18.1	20.5	4790	341	2717	229	55	1.3	8.1	7.53
T17	2.90	7.4	14.7	3600	270	620	124	42	1.33	7.7	7.82
T18	2.60	7.1	14.3	3500	260	610	116	39	2.16	6.3	7.32
T19	5.30	10.9	21.0	5000	400	720	110	53	1.99	11.8	7.25
T20	4.20	10.2	18.9	4300	350	780	111	52	1.54	12.4	7.25

Table C.1 continued.

ID	Ca (g kg^{-1})	Al (g kg^{-1})	Fe (g kg^{-1})	Mg (mg kg^{-1})	Mn (mg kg^{-1})	K (mg kg^{-1})	Na (mg kg^{-1})	Zn (mg kg^{-1})	Total C (g kg^{-1})	ASC (%)	pH
Reporoa Catchment											
R1	0.84	9.5	13.2	315	25	893	807	31	24	34.0	6.1
R2	0.82	4.4	6.8	539	70	487	474	47	1.73	14.1	NA
R3	1.08	5.7	5.7	603	187	735	708	35	1.23	11.2	NA
R4	2.12	4.5	11.6	989	672	840	650	68	3.82	16.8	NA
R5	1.00	3.7	5.6	515	81	748	750	24	1.28	11.7	NA
R6	1.05	5.6	9.3	607	97	650	632	28	14	31.0	4.9
R7	1.16	2.3	11.4	279	205	610	1620	16	30	20.0	6.5
R8	2.06	4.8	13.7	452	376	842	1828	29	17	20.0	5.7
R9	1.85	4.8	18.0	664	137	1032	1081	28	14	40.5	6.3
R10	1.29	4.4	8.2	576	58	506	756	25	4.09	31.9	NA
R11	1.50	7.0	5.6	528	64	667	1103	35	22.15	47.9	NA

Table C.2 Spearman correlations of the equilibrium phosphate concentration at net zero sorption (EPC_0) for the fresh (F) sediments, the difference in the sediment EPC_0 between freeze-dried (FD) and fresh sediments (i.e., dried minus fresh), and the difference in the sediment EPC_0 between air-dried (AD) and fresh sediments with the sediment physicochemical characteristics (see Table C.1); asterisks indicate the corresponding *p*-values ($p<0.05$, *; $p<0.01$, **) where the null hypothesis is no correlation between the variables. [a]Anion storage capacity; [b]Correlation not relevant; [c]The change in pH due to drying (i.e., pH of dried sediment minus pH of fresh sediment) for the given pre-treatment comparison

Variable	F	F vs. FD	F vs. AD
Ca	0.359	0.165	0.08
Al	0.655*	-0.464**	-0.457**
Fe	0.408	-0.311	-0.587*
Mg	0.503**	-0.428	-0.565*
Mn	0.293	-0.275	-0.563*
K	0.610*	-0.153	-0.053
Na	0.394	-0.078	0.296
Zn	0.614*	-0.368	-0.409
Total C	0.018	0.125	0.325
ASC[a]	-0.261	0.484**	0.251
pH	-0.135	-[b]	-
ΔpH, F vs. FD [c]	-	-0.276	-
ΔpH, F vs. AD	-	-	-0.054

Table C.3 For each fraction of sediment P in the fractionation scheme, Spearman correlations of the amount of P in the fresh (F) sediment P fraction, the difference in the sediment P fraction between freeze-dried (FD) and fresh sediments (i.e., dried minus fresh), and the difference in the sediment P fraction between air-dried (AD) and fresh sediments with the sediment physicochemical characteristics (see Table C.1); asterisks indicate the corresponding *p*-values ($p<0.05$, *; $p<0.01$, **) where the null hypothesis is no correlation between the variables. [a]Anion storage capacity; [b]Correlation not relevant; [c]The change in pH due to drying (i.e., pH of dried sediment minus pH of fresh sediment) for the given pre-treatment comparison

	NH_4Cl-P			**NaOH-RP**			**NaOH-URP**			**HCl-P**			**Residual P**		
Variable	F	F vs. FD	F vs. AD	F	F vs. FD	F vs. AD	F	F vs. FD	F vs. AD	F	F vs. FD	F vs. AD	F	F vs. FD	F vs. AD
Ca	0.265	0.095	-0.705**	-0.601**	0.300	0.444*	-0.268	0.164	0.127	0.749**	-0.187	-0.207	0.676**	0.120	-0.402*
Al	0.388*	0.111	-0.64**	-0.344	0.133	0.414*	-0.232	0.098	0.027	0.627**	0.0708	-0.0682	0.604**	0.155	-0.352
Fe	0.417*	0.013	-0.656**	-0.336	0.098	0.301	-0.4*	0.232	0.146	0.639**	0.00323	-0.139	0.606**	-0.050	-0.453*
Mg	0.488**	0.007	-0.765**	-0.457**	0.265	0.453*	-0.43*	0.281	-0.049	0.713**	-0.0242	-0.143	0.616**	-0.010	-0.469**
Mn	0.539**	-0.073	-0.504**	-0.192	0.027	0.134	-0.532**	0.386*	0.153	0.475**	0.0481	-0.164	0.36*	-0.148	-0.319
K	-0.003	-0.025	-0.177	0.070	-0.148	-0.005	0.101	-0.173	0.059	0.318	-0.269	-0.29	0.274	-0.078	-0.386*
Na	-0.609**	0.055	0.705**	0.707**	-0.501**	-0.611**	0.54**	-0.333	-0.163	-0.552**	-0.229	0.0385	-0.609**	-0.183	0.333
Zn	0.525**	-0.033	-0.602**	-0.262	0.057	0.258	-0.36*	0.251	-0.064	0.582**	-0.0198	-0.202	0.448*	0.010	-0.408*
Total C	-0.281	0.002	0.433*	0.546**	-0.456**	-0.316	0.373*	-0.192	0.064	-0.253	-0.125	0.0462	-0.343	0.047	0.216
ASC[a]	-0.531**	-0.128	0.669**	0.766**	-0.591**	-0.622**	0.505**	-0.276	-0.078	-0.401*	-0.235	-0.0167	-0.608**	-0.188	0.350
pH	0.083	-[b]	-	-0.627**	-	-	0.177	-	-	0.469*	-	-	0.399*	-	-
ΔpH, F vs. FD [c]	-	-0.107	-	-	0.182	-	-	0.143	-	-	-0.064	-	-	-0.013	-
ΔpH, F vs. AD	-	-	-0.110	-	-	0.327	-	-	-0.005	-	-	-0.0558	-	-	0.122

Appendix D

Supplementary Methods, Results, and Discussion for Chapter 4

D.1 Details on sediment phosphorus and iron fractionation

Bicarbonate-dithionite fractions

Since dithionite interferes with colorimetry (Lukkari et al. 2007), we treated the bicarbonate-dithionite (BD) extracts as follows. We first acidified BD extracts to maintain dissolved metals (0.8 mL of 1 M H2SO4 per 10 mL BD; Jensen and Thamdrup 1993), then allowed them to aerate overnight so the white, sulfur precipitant would settle out (Lukkari et al. 2007). Aliquots of the clear solution were digested as for total phosphorus (TP) analyses (see below).

Modified SEDEX fractionation to estimate authigenic Ca-P fractions

In addition to the main sediment P fractionation following Jan et al. (2015), which places more emphasis on metal oxides as the P pools of interest, we fractionated sediment P via a modified SEDEX procedure (Ruttenberg 1992) following Jensen et al. (1998). This modified SEDEX procedure complements the sediment P fractionation data by estimating an authigenic Ca-P fraction (i.e., authigenic apatite and $CaCO_3$-bound P, plus some leachable organic P) in addition to a detrital P fraction (i.e., primary mineral P) by extracting with solutions of increasing acidity.

Freeze-dried sediments were used as in the original SEDEX procedure (Ruttenberg 1992). We first extracted with 0.1 M NaOH first (two steps to avoid solution saturation, with a total extraction time of 20 h), which extracts sorbed P and P associated with metal oxides and organic matter (Hieltjes and Lijklema 1980; Jensen et al. 1998). We did not include a citrate-bicarbonate-dithionite step as in the original SEDEX scheme as citrate can complex with Ca and thus interfere with later Ca-focused fractions (Machesky et al. 2010). After wash steps with 1 M $MgCl_2$ (pH adjusted to 8; 2 h) and 0.5 M NaCl (15 min), sediments were extracted with a 1 M sodium acetate buffer (pH adjusted to 4 with acetic acid) for 6 h. We term this fraction as acetate-P. Following wash steps ($MgCl_2$ and NaCl washes), sediments were extracted with 0.5 M HCl for 24 h (termed SEDEX HCl-P); we expected this step to be similar to, but possibly lower than, the HCl-P step in the Jan et al. (2015) scheme depending on the amount of acetate-P present. Only the acetate-P and SEDEX HCl-P fractions were analyzed here.

Phosphorus colorimetry

We determined reactive P (RP) in the H_2O fraction via malachite-green (detection limit (DL) of 0.006 mg P L^{-1}; D'Angelo et al., 2001) and in the HCl fractions via molybdenum-blue (DL of ~0.02 mg P L-1; Murphy and Riley 1962). Digests were analyzed for TP with either colorimetric method depending on the required sensitivity. Acetate-P was determined by diluting at least ten-fold and neutralizing extracts prior to the molybdenum-blue method, keeping P concentrations below 150 µg P L^{-1} as suggested by Ruttenberg (1992). An external phosphate standard (1 mg PO_4 L^{-1}) was carried through analyses to ensure accurate and exchangeable measurements.

We initially measured RP in the NaOH steps as well with a molybdenum-blue method designed for alkaline extracts which does not hydrolyze organic P (He and Honeycutt 2005). However, we found substantial over-estimation (values greater than TP) afterwards. We suspect that the strong alkaline extractant dissolved substantial amounts of silicate minerals (Lindsay 1979; Sauer et al. 2006), which, unfortunately, would increase molybdenum-blue color development akin to reactive P (Nagul et al. 2015). The TP values were still valid since digestion removes silicate interference (Malá and Lagová 2014).

Iron colorimetry

Our colorimetric method for total iron (TFe) was a modification of the ferrozine method (Stookey 1970; Viollier et al. 2000). Since Fe oxidation state was irrelevant in the P fractionation scheme, we only measured TFe which required reduction of all Fe(III) to Fe(II) which reacts with ferrozine to produce a magenta color. Following Viollier et al. (2000): to 2.88 mL of the sample extract, we added 0.32 mL of ferrozine reagent and 0.6 mL of the reducing agent (1.4 M hydroxylamine hydrochloride in 2 M HCl); after reduction of all Fe(III) (see below), we added the ammonium acetate buffer (pH 9.5), swirled the vial (color appeared immediately), and read the absorbance at 562 nm. We allowed the mixture to reduce for 16 h rather than 10 min as in Viollier et al. (2000). Preliminary tests with Fe(III) standards ($FeCl_3$ in 0.01 M HCl) and Fe-spiked samples indicated incomplete reduction at greater Fe concentrations for times up to ~8 h. However, after 16 h under light conditions (to benefit from photochemical reduction; Anastácio et al. 2008), standard curves were linear up to 75 µM Fe, replicable, and stable for at least several hours (Stookey 1970). The method detection limit with a 1 cm light path was approximately 0.3 µM Fe.

Quality control and checks

To ensure replicable results, we included an internal reference sediment in all batches of P fractionation and subsequent analyses. The internal reference was a freeze-dried, relatively homogenous floodplain sediment with a sandy texture. Notably, the internal reference alerted us to the issue of an expired

dithionite chemical used early in the study: the batches affected were immediately repeated with a newer chemical.

We measured Fe and P in blanks for each fractionation step to account for possible contamination. Here, either zero or very low concentrations were measured in extractant/digest blanks (e.g., <10 µg P L^{-1} and <0.2 mg Fe L^{-1}; typically, 2 or more orders of magnitude lower than the samples) and the data were corrected for these blanks accordingly.

D.2 Linear models

To aid in our objective of relating in-stream DRP to sediment chemistry, we tested simple linear models on (1) DRP and (2) for sorption metrics (anion storage capacity (ASC) only, as Bache-Williams index had similar results). We use these predictive models to generate causal hypotheses to test in future studies (Shmueli 2010).

For DRP, we fitted linear models primarily with H_2O-P, BD-I P, and ASC, including geology as a grouping variable. We applied a similar approach for modelling ASC, but with pools of Fe (BD-I, BD-II, and total Fe) as the primary variables of interest. Residual checks on initial model fits indicated problems with bias, heteroskedasticity, and points with high leverage. Therefore, rather than ordinary least-squares regression, we applied robust regression (MASS package; Venables and Ripley 2002) with Huber's weighting scheme. Unfortunately, this method of estimating a linear model precludes estimates of standard errors about the fit; only the model fits themselves are shown in the figures.

While formal model comparison tests (e.g., F-tests) were not applicable, we compared model performance using Akaike's AIC and root mean square error (RMSE). Although such simple linear models of the complex cycling of P in streams will have limited predictive ability, we utilized these models for discussion purposes.

D.3 Supplementary results for solution equilibria

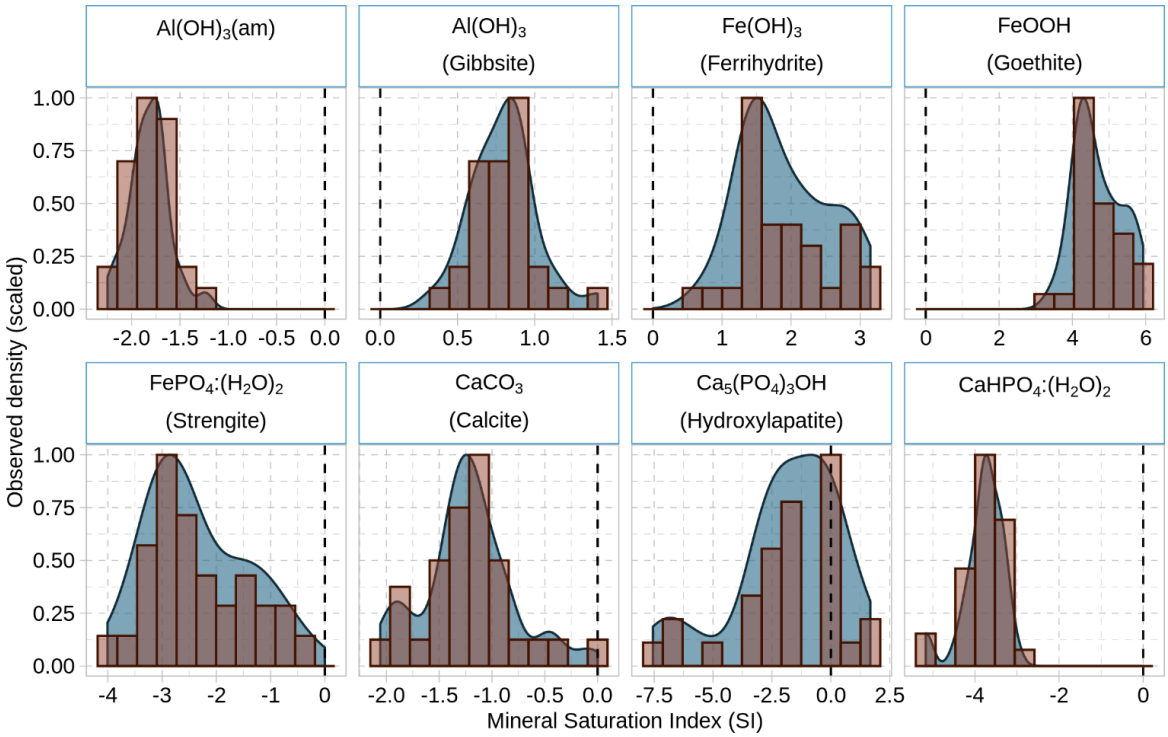

Figure D.1 Distributions of select mineral saturation indices for the stream samples (n=31), where positive (negative) saturation index (SI) indicates the thermodynamic potential for the mineral to precipitate (dissolve). The frequency distributions are displayed as normalized densities. SI=0 is indicated on each sub-plot with a dashed vertical line. The solid phase chemical formulae are provided as given in the MINTEQ.v4 database.

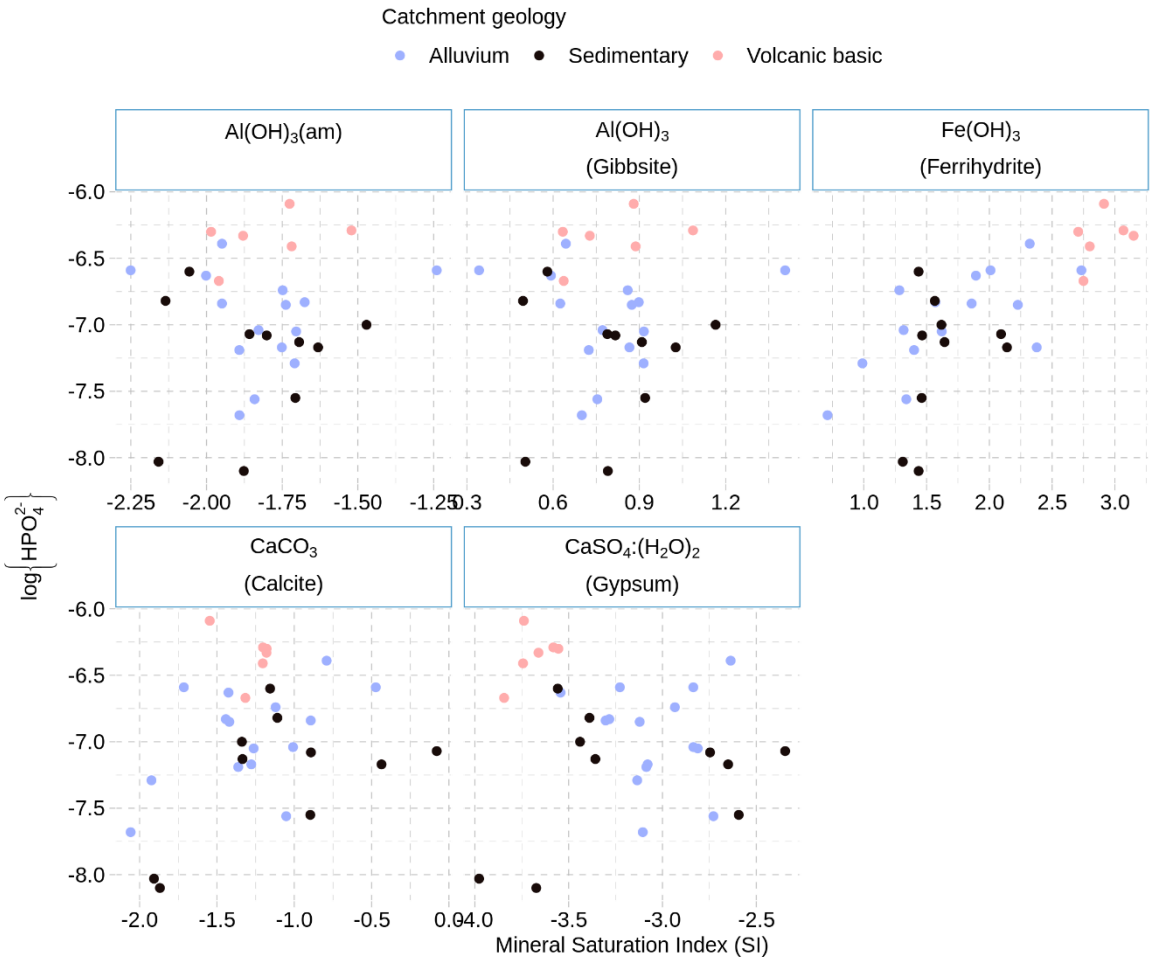

Figure D.2 Log-activity of HPO42- as a function of select mineral saturation indices (SIs). Mineral formulas are shown as given in the MINTEQA2 v4 database. Note that ferrihydrite activity is likely inflated due to colloidal Fe species.

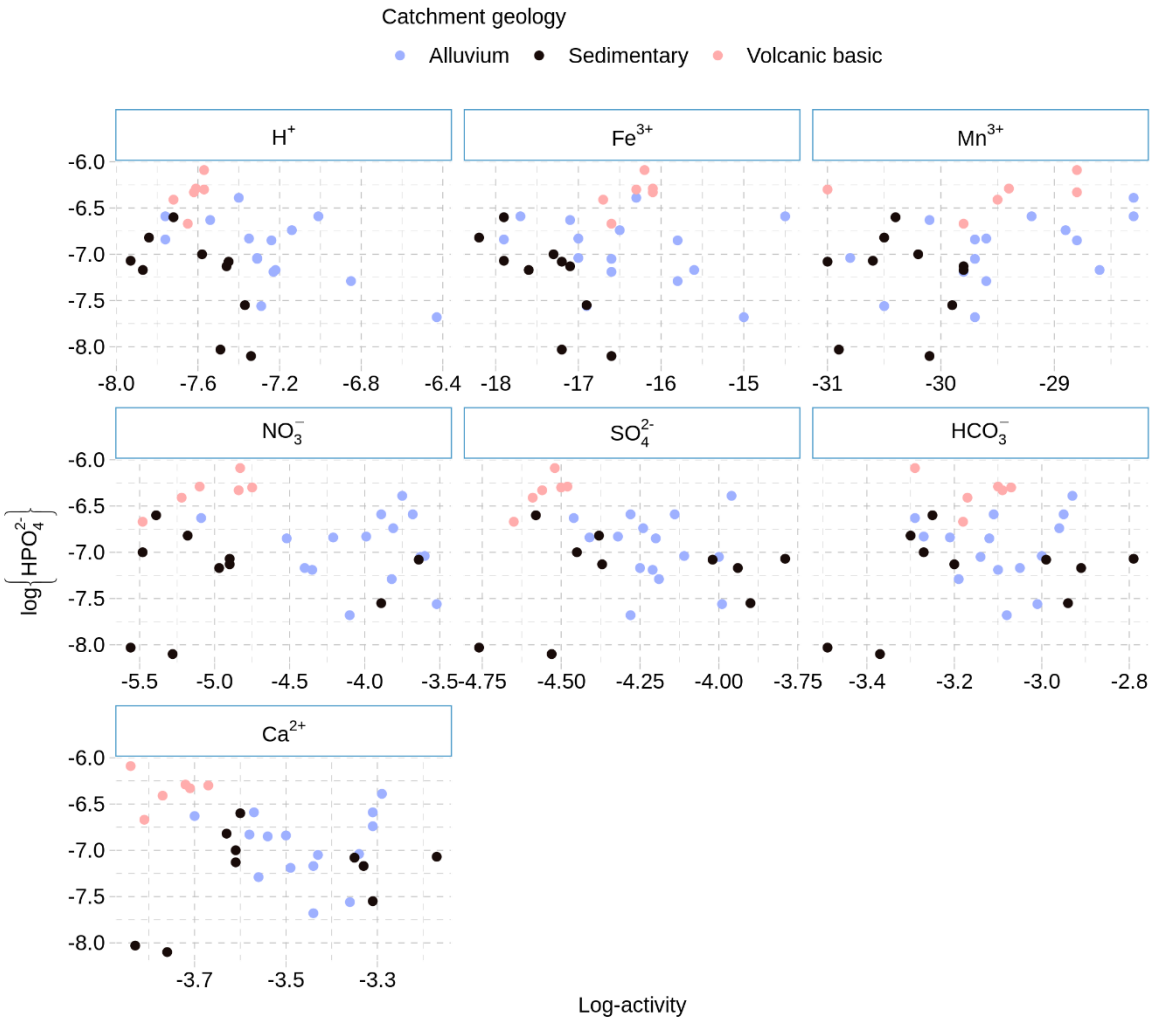

Figure D.3 Log-activity of HPO42- as a function of select ion log-activities as modelled by PHREEQC.

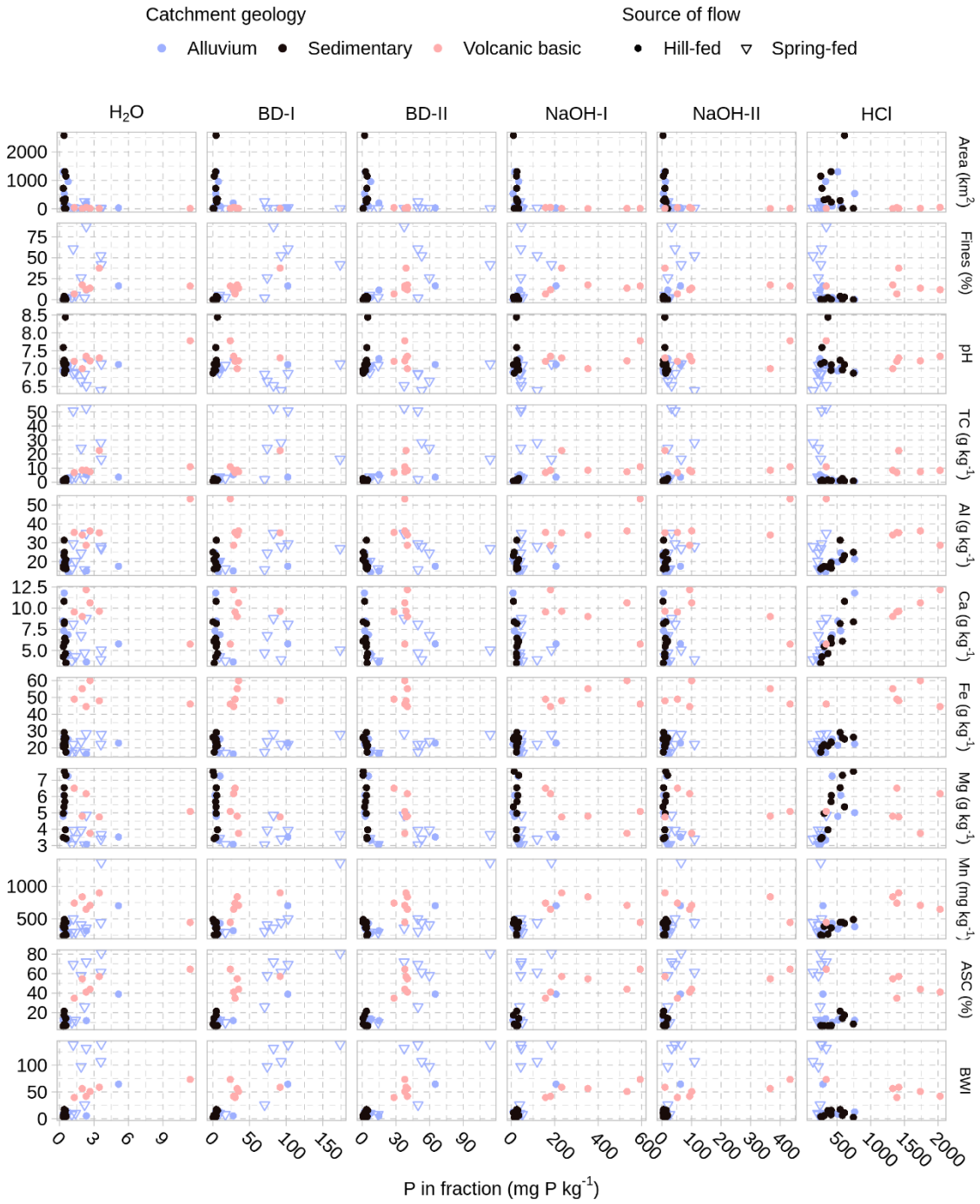

Figure D.4 Scatter plots of sediment P fractions and select sediment physicochemical variables; while P fractions are in mg P kg^{-1}, each other variable has units given in the label (except Bache-Williams Index (BWI), for which we refer the reader to the main text). Note the change in scales for each variable and that the physicochemical variables are plotted on the *y*-axes out of consideration of space.

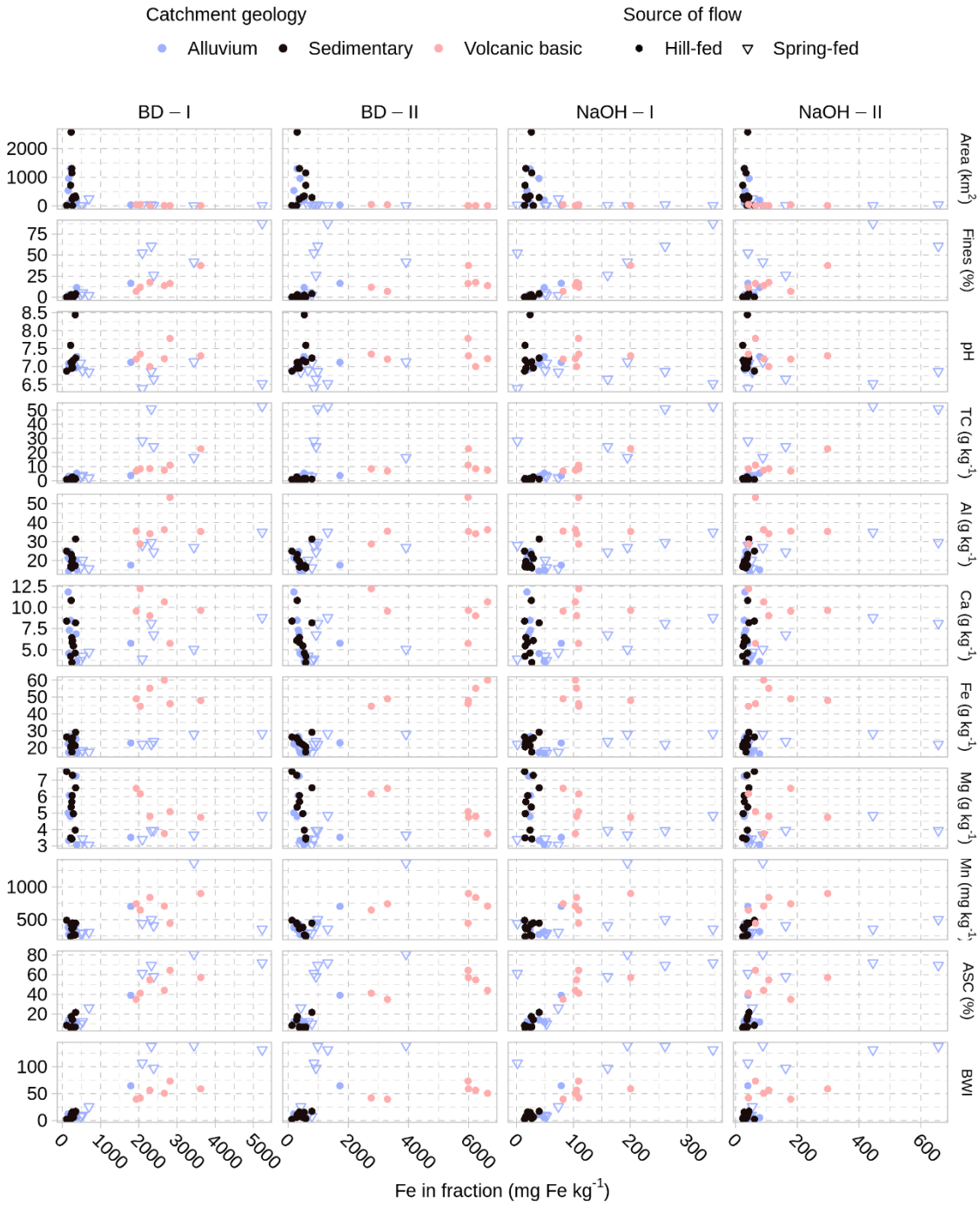

Figure D.5 Scatter plots of Fe in sediment P fractions and select sediment physicochemical variables; while Fe fractions are in mg Fe kg^{-1}, each other variable has units given in the label (except Bache-Williams Index (BWI), for which we refer the reader to the main text). Note the change in scales for each variable and that the physicochemical variables are plotted on the *y*-axes out of consideration of space.

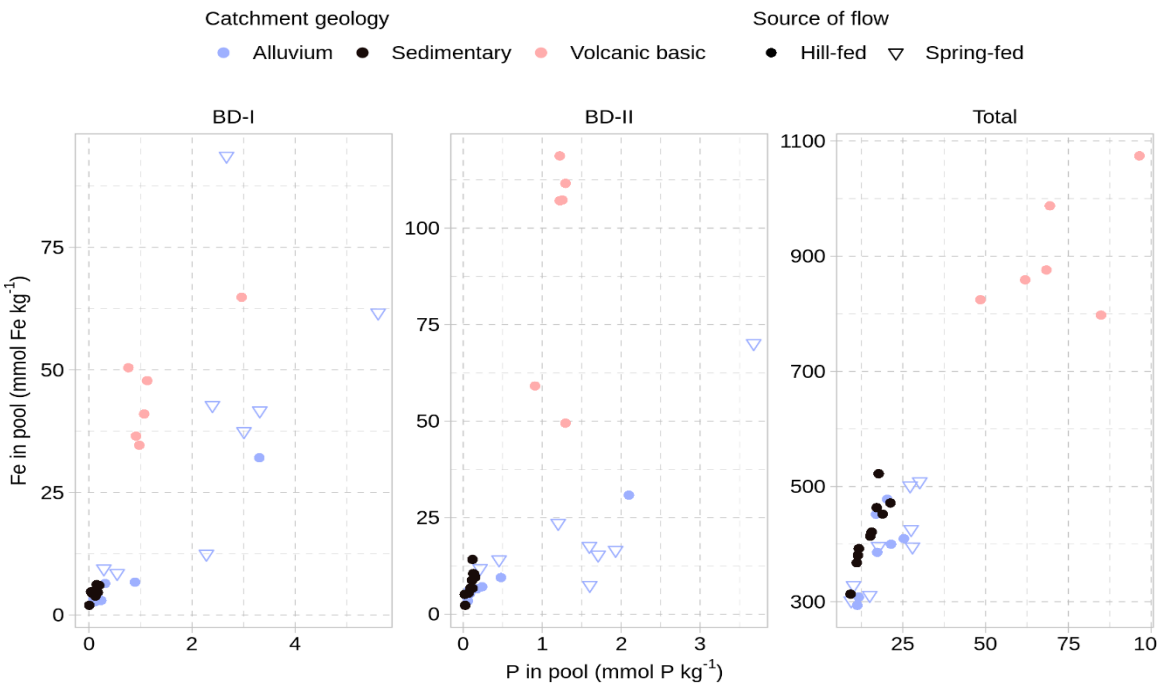

Figure D.6 Sediment Fe and P content for the bicarbonate-dithionite (BD) extractions and total content; units are on molar basis to facilitate comparison of Fe:P ratios.

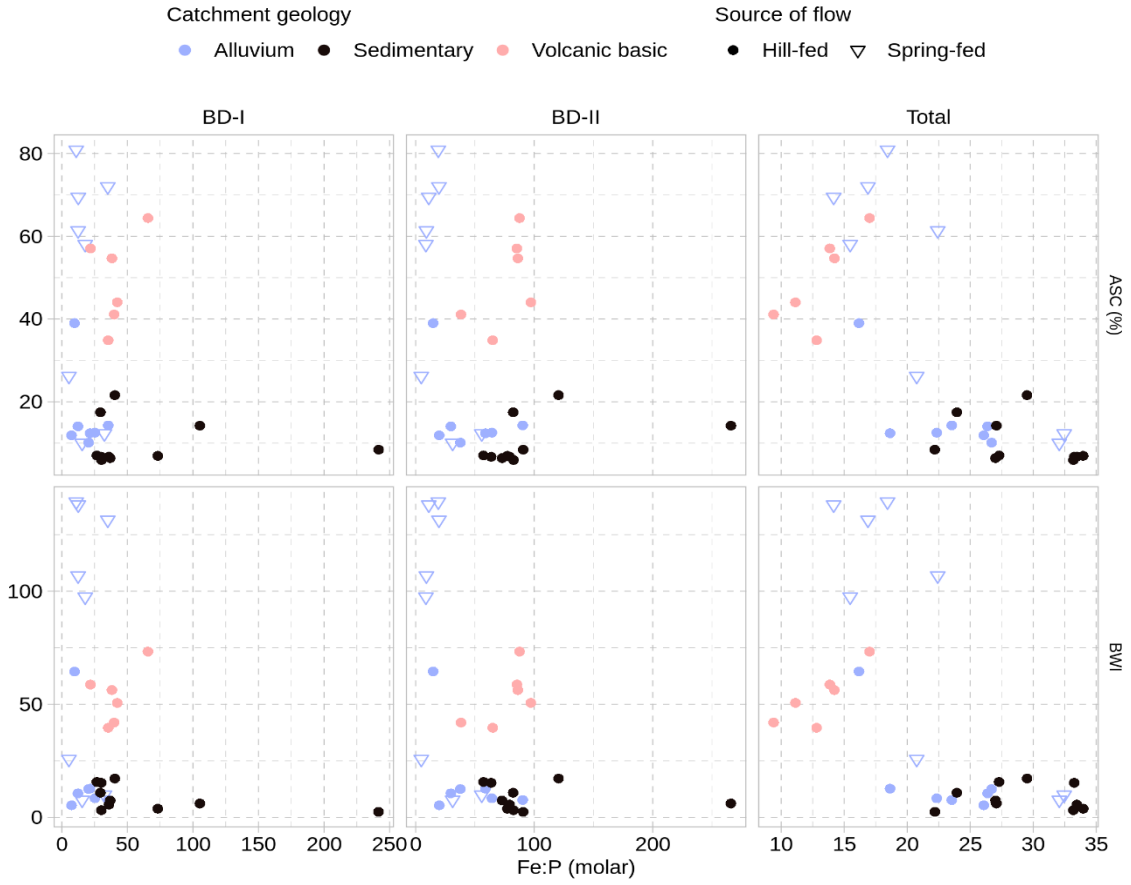

Figure D.7 Sediment sorption metrics (anion storage capacity (ASC), expressed as a %, and Bache-Williams Index (BWI), whose units are given in the main text) plotted against sediment molar Fe:P ratios for the bicarbonate-dithionite (BD) fractions and total content.

Table D.1 Stream study sites and their characteristics generated from the River Environment Classification database grouped according to the three main geology classes surveyed. The source of flow (i.e., topography), geology (derived from New Zealand Land Resources Inventory 'toprock' geology data), and land-cover (1997) classifications are determined through the pre-dominant characteristics of the catchment (Snelder et al. 2010). Strahler stream order is also given. The sites are located at monitoring stations (https://www.lawa.org.nz/) except for the two Craigieburn sites (†).

Site name	Catchment area (km^2)	Source of flow / Spring-fed	Land-cover (1997)	Stream order
Alluvium				
Taranaki @ Preeces	0.968	Low-elevation / Spring-fed	Pastoral	1
Waikuku	4.07	Low-elevation / Spring-fed	Pastoral	2
Taranaki @ Greesons	8.71	Low-elevation / Spring-fed	Urban	2
Knights @ Saby's Rd	12.6	Low-elevation / Spring-fed	Pastoral	3
Halswell @ Akaroa Br	15.8	Low-elevation / Spring-fed	Pastoral	3
Saltwater Creek	24.1	Low-elevation / Spring-fed	Pastoral	4
Waianiwaniwa	31.5	Low-elevation	Pastoral	4
L-II Stream @ Pannet Br	38.1	Low-elevation / Spring-fed	Pastoral	3
Hawkins River	92.8	Hill	Pastoral	4
N Ashburton @ Digby Br	98.2	Low-elevation	Pastoral	4
Cust @ Skewbridge	203	Low-elevation	Pastoral	5
Halswell @ McCartney	251	Low-elevation / Spring-fed	Pastoral	5
S Ashburton @ Quarry Rd	535	Hill	Tussock	6
Selwyn River @ Coes Ford	958	Low-elevation	Pastoral	5
S Ashburton @ Hills Rd	1300	Hill	Pastoral	6
Sedimentary (Hard and Soft)				
Craigieburn @ Cave Stream†	5.1	Hill	Indigenous forest	2
Craigieburn @ Dracophyllum track†	14	Mountain	Bare ground	3
Pahau @ Dalzell's farm	235	Hill	Pastoral	5
N Ashburton @ SH 72	291	Mountain	Tussock	5
Waitohi	315	Hill	Pastoral	5
Waipara @ Laidmore Rd	345	Hill	Pastoral	6
Waipara @ SH1	719	Low-elevation	Pastoral	6
Ashley River @ SH1	1150	Hill	Pastoral	7
Hurunui @ SH7	1310	Hill	Indigenous forest	6
Rakaia @ SH 77	2580	Glacial-Mountain	Bare ground	7
Volcanic Basic				
French Farm	6.65	Low-elevation	Pastoral	3
Wainui	10.3	Low-elevation	Pastoral	3
Barry's Bay	11.2	Low-elevation	Pastoral	3
Takamatua @ SH75	12.6	Low-elevation	Pastoral	3
Kaituna	40.7	Low-elevation	Pastoral	4
Okana	48.4	Low-elevation	Pastoral	4

Table D.2 Summary of best-fit robust linear models for DRP (µg P L-1) using catchment geology, pools of sediment P (H_2O and the first bicarbonate-dithionite (BD-I) P fractions), and anion storage capacity (ASC); data from 'spring' sites were excluded from this analysis (see main text; here, *n*=23). A model with only geology is included for comparison. DF is model degrees of freedom, RMSE is root mean square error, and AIC is Akaike's 'An Information Criterion'.

Terms in linear model	DF	RMSE (µg P L^{-1})	Akaike's AIC
Geology	3	6.73	161.0
H_2O-P	2	5.72	151.5
Geology, H_2O-P	4	3.93	138.2
Geology, H_2O-P, BD-I P	5	3.84	139.2
ASC	2	4.44	139.9
ASC, H_2O-P	3	4.38	131.9

Table D.3 Summary of best-fit robust linear models for anion storage capacity (%) using catchment geology and pools of sediment Fe in the bicarbonate-dithionite extractions (BD-I, BD-II) and total digest as predictors (*n*=31); a model with only geology is included for comparison. DF is model degrees of freedom, RMSE is root mean square error, and AIC is Akaike's 'An Information Criterion'.

Terms in linear model	DF	RMSE (%)	Akaike's AIC
Geology	3	19.3	279.5
BD-I Fe	2	9.17	231.4
BD-II Fe	2	19.0	276.5
Acetate-P	2	13.7	256.0
BD-I Fe, Acetate-P	3	8.45	228.3
Geology, BD-II Fe	4	15.8	269.2
Total Fe	2	21.1	283.1
Geology, Total Fe	3	17.7	276.2

D.4 Supplementary Discussion: Solution geochemical equilibria

Phosphate-minerals rarely had the thermodynamic potential to form in these waters. The most stable phosphate mineral, hydroxylapatite, showed some cases of near-saturation (SI ≈ 0; Figure 4.2). However, empirical research has suggested that the required supersaturation for hydroxylapatite to significantly precipitate from solution is much higher (reported SI from 3 to 10; House 1999; Plant and House 2002; Sø et al. 2011). Therefore, hydroxylapatite and other phosphate-mineral precipitation seems an unlikely mechanism for P removal in these streams. Streams with greater Ca concentrations (e.g., >100 mg Ca L^{-1}) and pH (>8) will likely be more conducive for Ca-P mineral precipitation (Diaz et al. 1994). It is striking that some streams approach, but do not significantly extend beyond, the hydroxylapatite solubility curve, which could suggest a role for hydroxylapatite equilibrium. However, stream solution chemistry is more complex than ideal solutions modelled by PHREEQC, where even the solubility constant for hydroxylapatite is subject to considerable uncertainty (Golterman 2004). In addition, other phosphate-minerals (e.g., various Al and Fe based phosphate-minerals) were unlikely to contribute to phosphate activity in the water column, but may be more important in some subsurface environments (Rothe et al. 2014).

Since, calcite co-precipitates with phosphate (House 2003; Golterman 2004), periods of calcite precipitation in streams may provide an opportunity for phosphate removal, particularly in low-phosphate systems (Plant and House 2002; Machesky et al. 2010; Sø et al. 2011). House (1999) suggested that calcite precipitation in streams becomes significant near SI ≈ 1. In the present study, however, we only observed negative SI's for calcite (Figure D.1). Our study design was more likely to capture the upper extent of calcite SI variability since: 1) calcite saturation peaks during the day when photosynthesis depletes CO_2(g) and increases pH (Nimick et al. 2011; Stets et al. 2017) and 2) calcite solubility decreases with greater temperatures during the day (Stumm and Morgan 1996). It is likely that our study streams were not alkaline enough for significant calcite interactions (here, median alkalinities of 38.2 to 42.1, maximum of 88.8 mg L^{-1} as $CaCO_3$; maximum pH of 7.93; Table 4.2) since streams that reach a calcite SI ≥ 1 typically have alkalinity >100 mg L^{-1} as $CaCO_3$ and sustain pH > 8 (Neal et al. 2002; Nimick et al. 2011; Corman et al. 2015; Stets et al. 2017).

It is important to note the supersaturation of ferrihydrite and its relationship with HPO_4^{2-} activity (Figure D.2). Ferric iron is largely insoluble in most stream conditions (oxygenated water and pH near or above neutral) and Fe(II) would presumably be associated only with reducing zones within the stream corridor (Stumm and Sulzberger 1992), thus positive ferrihydrite SI seems implausible. Fox (1988) explained the problem that, in most streams, dissolved Fe is overestimated because common filtration methods (i.e., 0.45 µm filters) fail to remove colloidal Fe species (van der Grift et al. 2014; Baken et al. 2016a). Our apparent supersaturation with respect to ferrihydrite is within the range of previously observed over-estimates (SI up to 5 in most cases; Fox 1988). However, the positive relation between ferrihydrite SI and DRP may indirectly point towards evidence of colloidal Fe species carrying sorbed phosphate and should be investigated in future research.

Appendix E

Supplementary Methods and Results for Chapter 5

E.1 Stream sites information

Table E.1 Study streams across Banks Peninsula, New Zealand. ID is an arbitrary identifier used throughout the rest of the text, order refers to stream order, and area is catchment area. Land uses within the catchments were obtained from LCDB v5.0 (https://lris.scinfo.org.nz/), with some classes aggregated for brevity; other classes not shown were all <5% of the total area.

Stream site	**ID**	**Order**	**Area (km^2)**	**Land use (%)**					
				Grassland, intensive	Grassland, unintensive	Forest, exotic	Forest, native or unintensive	Shrub/Scrub	Urban
Narbey Stream	BP1	3	8.6	8.9	3.7	4.1	35.2	48.1	0.0
Le Bons Stream	BP2	3	18.4	67.5	0.0	4.9	18.4	9.0	0.1
Aylmers Stream	BP3	2	3.7	52.7	4.6	7.5	18.3	11.8	5.0
Balguerie Stream	BP4	2	4.4	35.3	12.4	5.2	17.5	23.0	6.0
Takamatua Stream	BP5	3	12.8	53.1	0.8	5.6	19.9	18.9	0.7
Stream at Little Akaloa Rd	BP6	3	12.9	73.4	0.7	2.6	12.5	10.8	0.0
Barry's Bay Stream	BP7	3	11.5	82.1	0.0	1.7	8.5	6.7	0.0
Stream at Wainui Valley Rd	BP8	3	10.0	54.3	5.7	6.4	16.7	13.5	0.3
Stream at French Farm Valley Rd	BP9	3	8.0	77.2	0.0	2.4	11.5	8.6	0.0
Okana River	BP10	4	50.9	70.3	3.4	7.2	11.9	6.8	0.5
Okuti River	BP11	3	26.2	56.1	7.2	10.6	10.8	15.2	0.0
Kaituna River	BP12	4	33.1	55.2	14.5	2.6	16.6	11.0	0.0

E.2 Nonlinear mixed effects modeling for the nutrient addition experiment

The objective of this experiment was to determine whether the additions of bioavailable C and/or N to the solutions used for EPC_0 would decrease EPC_0 via increased biotic demand for P. Through the factorial experimental design (nutrient treatments of 'None', +N, +C, and C+N for all 12 sediments), a simple approach to test the null hypothesis (no change in EPC_0 across the nutrient treatments) would be a two-way ANOVA on the resultant EPC_0 (n=12). However, this analysis would not take into account the imprecision behind the EPC_0 (which is not equal across sediments and treatments) and so not fully leverage all the information in the raw dataset (here, nominally 9 sorption points per sediment after discarding sorption points for the 250 µg P L^{-1} treatment). Therefore, we tested for nutrient effects on EPC_0 while accounting for all measurements constituting each EPC_0.

To do this, we used nonlinear mixed effects modelling (R package 'nlme' (Pinheiro et al. 2020)) to directly test for changes in EPC_0 due to nutrient treatment. The linear sorption model was re-expressed in a nonlinear form to make EPC_0 a parameter to estimate:

$$q = \beta(C_i - EPC_0) \qquad (1)$$

where q is sorption (mg P kg^{-1}), C_i is the initial concentration (µg P L^{-1}), β is the linear slope (mg P kg^{-1} per µg P L^{-1}) and EPC_0 is the x-intercept corresponding to a q of 0 mg P kg^{-1}. We use initial rather than final concentrations as the more appropriate independent variable (for more details, see supplementary information in Simpson et al. 2019). By modelling EPC_0 directly as a parameter, we can employ fixed effects to account for the variance in EPC_0 due to nutrient treatment. Additionally, since EPC_0 naturally varies by sediment, we used random effects for the EPC_0 parameter, where deviations in overall mean EPC_0 occurred by the 12 sediments.

We fitted increasingly complex models, where the two parameters (β and EPC_0) were allowed to vary with nutrient treatment (fixed effects) and to vary by individual sediment (random effect). We tested models through the likelihood ratio test and by comparing AIC values (Pinheiro and Bates 2000) and further judged model fits through fixed effect size (and 95% confidence intervals) and diagnostic plots.

A full model (i.e., an individual fit of eq. 1 for every sediment × nutrient treatment dataset (n=*48*), resulting in 96 model parameters) suggested that β varies little across the whole data; this was confirmed when modeling with mixed-effects by stripping away sediment and nutrient treatment effects for β with no loss in model quality. EPC_0 varied by sediment and no higher order random effects (i.e., nutrient treatment nested within sediment for EPC_0) were necessary. With random effects only for EPC_0, a model with EPC_0 fixed effects for +C (p=0.0698), +N (p=0.0059), and their interaction (p=0.6181) suggested significant effects only for N addition. Indeed, this model suggested no improvement over a simpler model without the +C and interaction terms for EPC_0 (p=0.102). In summary, our final model included random effects by sediment for EPC_0 and only N addition as an added fixed effect (Table 5.4). The addition of N to the EPC_0 solutions suggested a 0.56 µg P L^{-1} drop in measured EPC_0.

E.3 Supporting Results

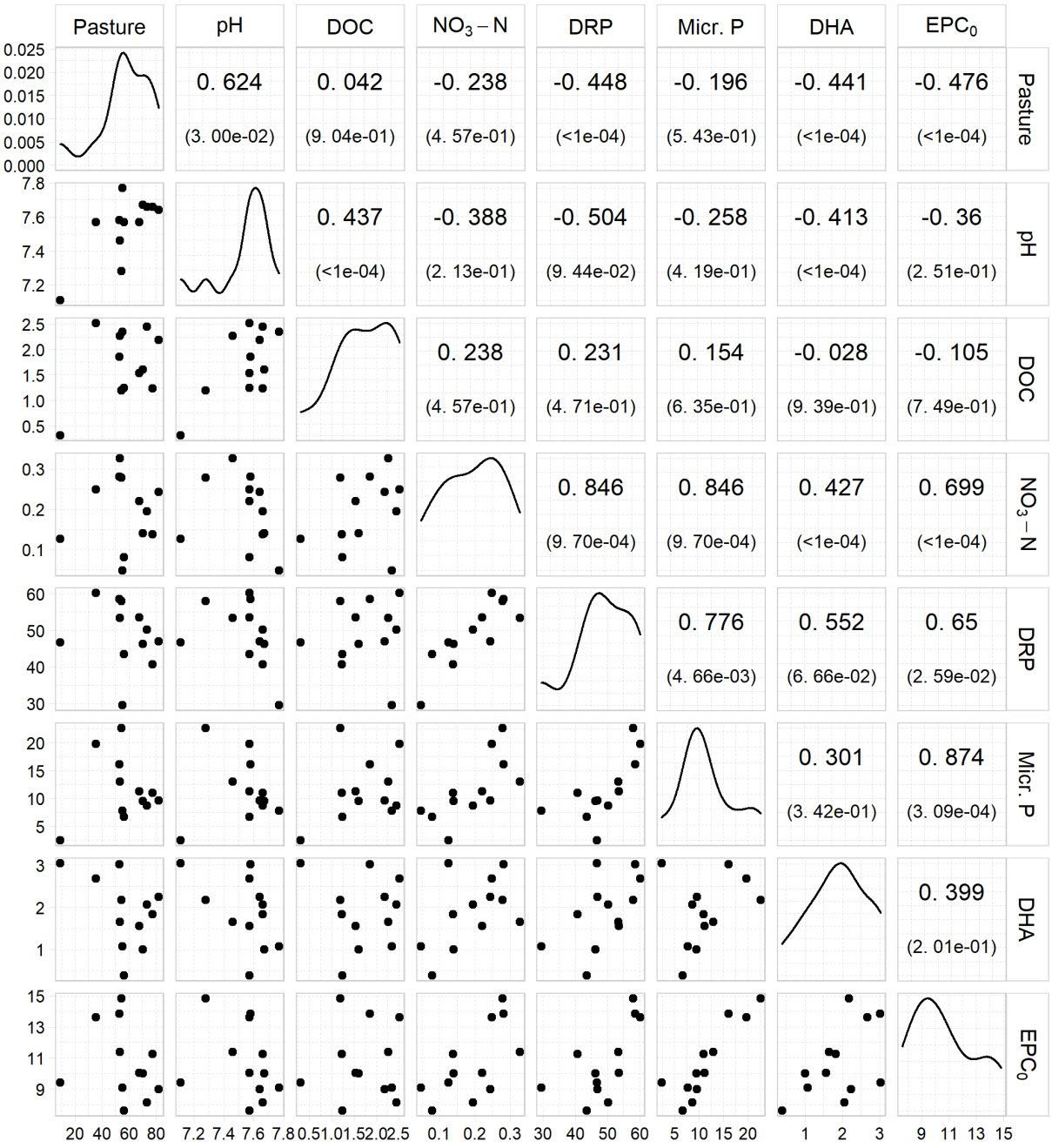

Figure E.1 Correlation matrix for select stream water variables, pasture land use, and primary sediment variables. 'Micr.' is the abbreviation for 'microbial biomass'; see main text for other abbreviations. Units are not shown for brevity, but are (in order shown, from pasture to EPC_0): % area, S.U., mg C L^{-1}, mg N L^{-1}, µg P L^{-1}, mg P kg^{-1}, mg TPF kg^{-1} hr^{-1}, and µg P L^{-1}. The upper part of the matrix reports Spearman correlations (*p*-values in parentheses). The diagonal is a kernel density plot for the variable.

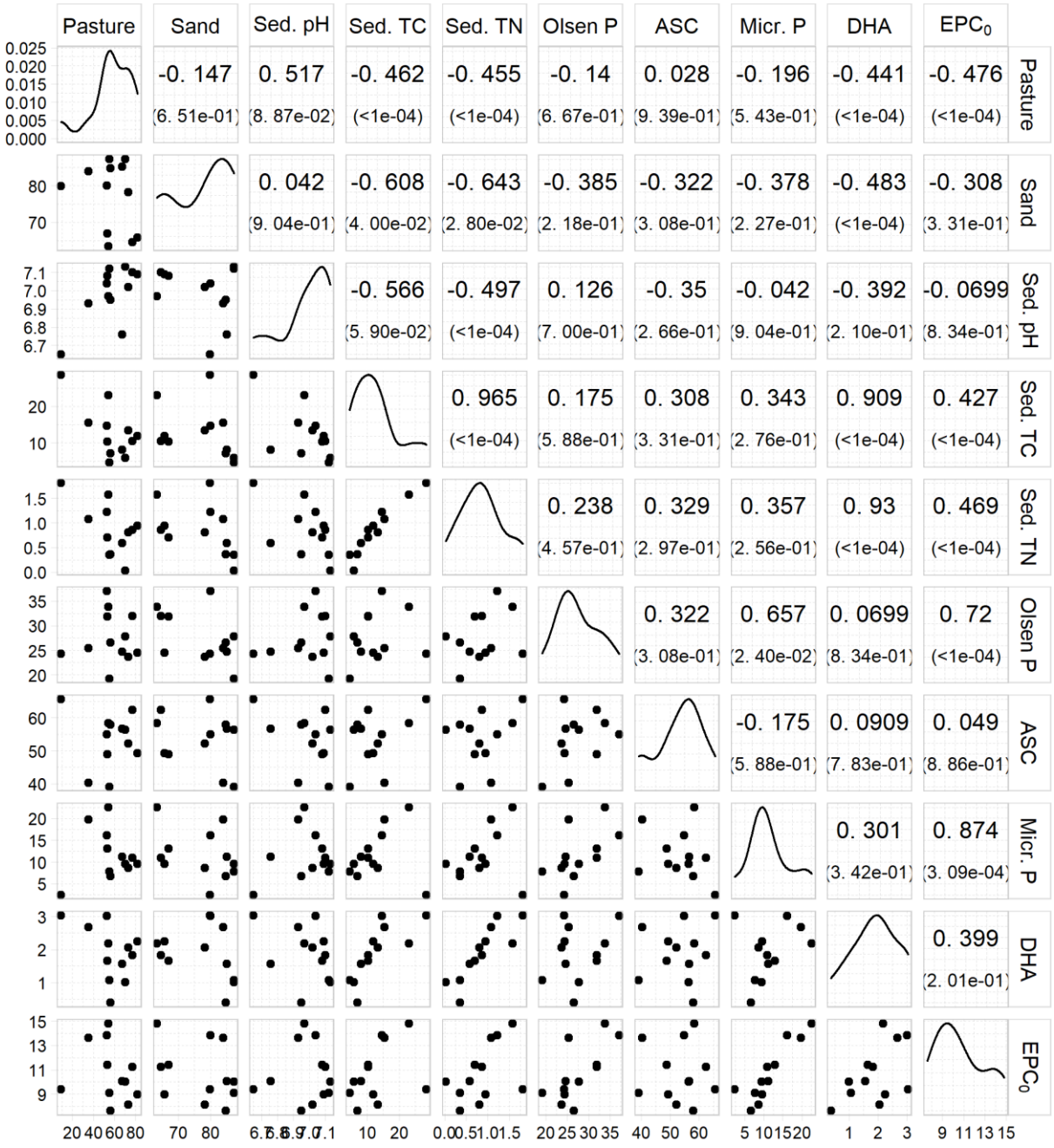

Figure E.2 Correlation matrix for select sediment physicochemical variables, pasture land use, and primary sediment variables. 'Micr.' and 'Sed.' are the abbreviations for 'microbial biomass' and 'sediment'; see main text for other abbreviations. Units are not shown for brevity, but are (in order shown, from pasture to EPC_0): % area, % particles, S.U., g C kg^{-1}, g N kg^{-1}, mg P kg^{-1}, % P retention, mg P kg^{-1}, mg TPF kg^{-1} hr^{-1}, and µg P L^{-1}. The upper part of the matrix reports Spearman correlations (*p*-values in parentheses). The diagonal is a kernel density plot for the variable.

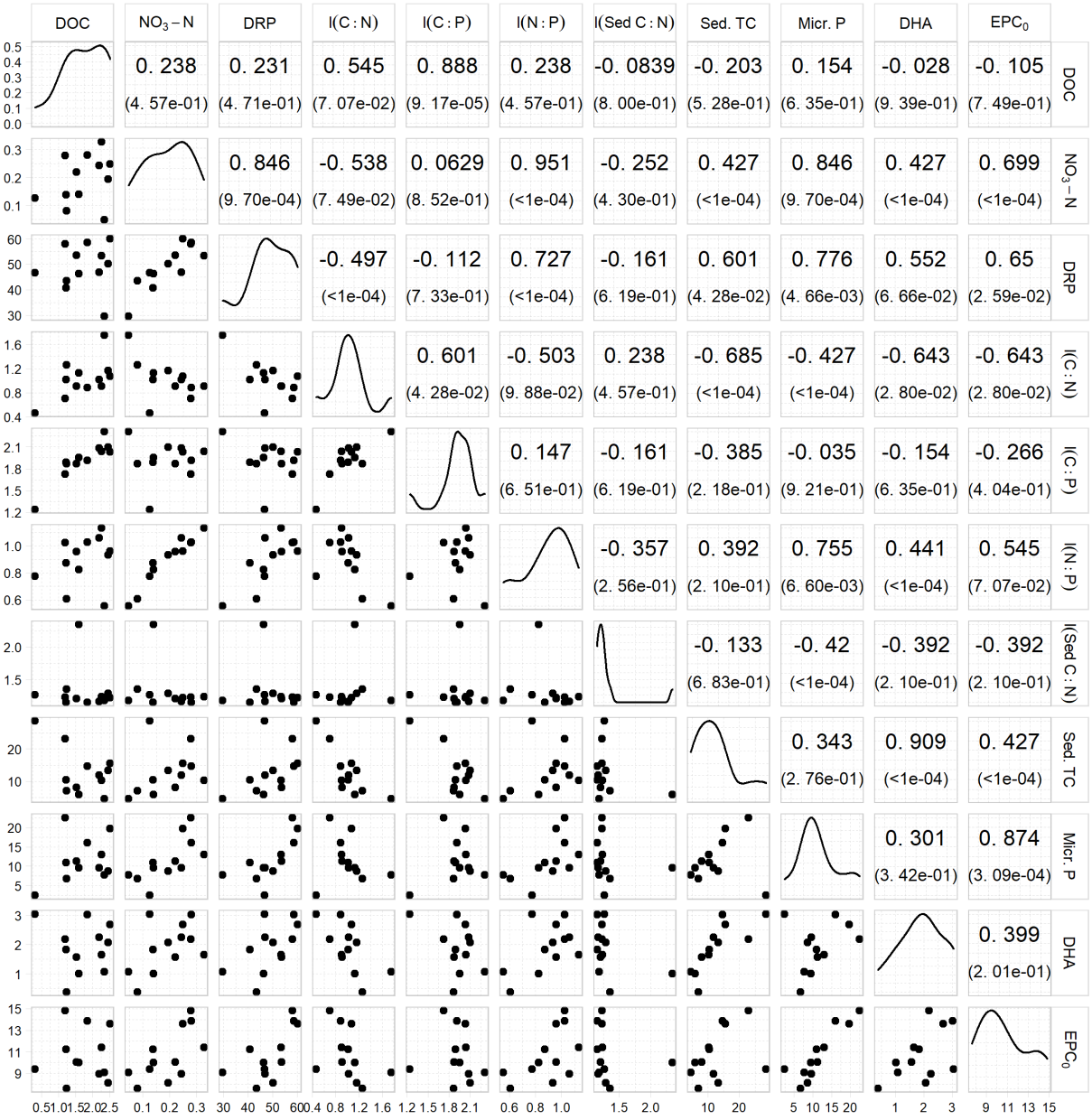

Figure E.3 Correlation matrix for stream and sediment C/nutrient ratios, select stream/sediment physicochemical variables, and primary sediment variables. 'Micr.' and 'Sed.' are the abbreviations for 'microbial biomass' and 'sediment'; see main text for other abbreviations. All elemental ratios are on molar basis and are $\log_{10}$ transformed (referred to with 'l()') for visual clarity (note: transformations have no bearing on the Spearman correlation). Units are not shown for brevity, see captions for Figures E1 and E2 for details. The upper part of the matrix reports Spearman correlations (*p*-values in parentheses). The diagonal is a kernel density plot for the variable.

References

Acuña V, Wolf A, Uehlinger U, Tockner K (2008) Temperature dependence of stream benthic respiration in an Alpine river network under global warming. Freshwater Biology 53:2076–2088. https://doi.org/10.1111/j.1365-2427.2008.02028.x

Agudelo SC, Nelson NO, Barnes PL, et al (2011) Phosphorus adsorption and desorption potential of stream sediments and field soils in agricultural watersheds. Journal of Environment Quality 40:144. https://doi.org/10.2134/jeq2010.0153

Anastácio AS, Harris B, Yoo HI, et al (2008) Limitations of the ferrozine method for quantitative assay of mineral systems for ferrous and total iron. Geochim Cosmochim Acta 72:5001–5008. https://doi.org/10.1016/j.gca.2008.07.009

APHA (2005) Standard Methods for the Examination of Water and Wastewater, 21st Edition. American Public Health Association, USA

Appelo CAJ, Postma D (2005) Geochemistry, groundwater and pollution, 2nd edn. A.A. Balkema Publishers, Leiden, The Netherlands

Arai Y, Sparks DL (2007) Phosphate Reaction Dynamics in Soils and Soil Components: A Multiscale Approach. Advances in Agronomy 94:135–179. https://doi.org/10.1016/S0065-2113(06)94003-6

Arai Y, Sparks DL (2001) ATR–FTIR Spectroscopic Investigation on Phosphate Adsorption Mechanisms at the Ferrihydrite–Water Interface. Journal of Colloid and Interface Science 241:317–326. https://doi.org/10.1006/jcis.2001.7773

Ardón M, Zeglin LH, Utz RM, et al (2021) Experimental nitrogen and phosphorus enrichment stimulates multiple trophic levels of algal and detrital-based food webs: a global meta-analysis from streams and rivers. Biological Reviews 96:692–715. https://doi.org/10.1111/brv.12673

Attygalla NW, Baldwin DS, Silvester E, et al (2016) The severity of sediment desiccation affects the adsorption characteristics and speciation of phosphorus. Environmental Science: Processes & Impacts 18:64–71. https://doi.org/10.1039/C5EM00523J

Aubeneau AF, Hanrahan B, Bolster D, Tank JL (2014) Substrate size and heterogeneity control anomalous transport in small streams. Geophys Res Lett 41:8335–8341. https://doi.org/10.1002/2014GL061838

Audette Y, O'Halloran IP, Nowell PM, et al (2018) Speciation of Phosphorus from Agricultural Muck Soils to Stream and Lake Sediments. Journal of Environment Quality 47:884. https://doi.org/10.2134/jeq2018.02.0068

Azizian M, Boano F, Cook PLM, et al (2017) Ambient groundwater flow diminishes nitrate processing in the hyporheic zone of streams. Water Resources Research 53:3941–3967. https://doi.org/10.1002/2016WR020048

Bache B, Williams E (1971) A phosphate sorption index for soil. Eur J Soil Sci 22:289–301. https://doi.org/10.1111/j.1365-2389.1971.tb01617.x

Baken S, Moens C, van der Grift B, Smolders E (2016a) Phosphate binding by natural iron-rich colloids in streams. Water Res 98:326–333. https://doi.org/10.1016/j.watres.2016.04.032

Baken S, Regelink IC, Comans RNJ, et al (2016b) Iron-rich colloids as carriers of phosphorus in streams: A field-flow fractionation study. Water Res 99:83–90. https://doi.org/10.1016/j.watres.2016.04.060

Baker MA, Webster JR (2017) Chapter 30 - Conservative and Reactive Solute Dynamics. In: Lamberti GA, Hauer FR (eds) Methods in Stream Ecology (Third Edition). Academic Press, pp 129–145

Baldwin DS (1996) Effects of exposure to air and subsequent drying on the phosphate sorption characteristics of sediments from a eutrophic reservoir. Limnology and Oceanography 41:1725–1732. https://doi.org/10.4319/lo.1996.41.8.1725

Baldwin DS (2013) Organic phosphorus in the aquatic environment. Environ Chem 10:439–454. https://doi.org/10.1071/EN13151

Ball DF (1964) Loss-on-Ignition as an Estimate of Organic Matter and Organic Carbon in Non-Calcareous Soils. Journal of Soil Science 15:84–92. https://doi.org/10.1111/j.1365-2389.1964.tb00247.x

Barbanti A, Bergamini MC, Frascari F, et al (1994) Critical Aspects of Sedimentary Phosphorus Chemical Fractionation. Journal of Environmental Quality 23:1093–1102. https://doi.org/10.2134/jeq1994.00472425002300050035x

Barrow NJ (1983a) A mechanistic model for describing the sorption and desorption of phosphate by soil. Journal of Soil Science 34:733–750. https://doi.org/10.1111/j.1365-2389.1983.tb01068.x

Barrow NJ (1983b) On the reversibility of phosphate sorption by soils. European Journal of Soil Science 34:751–758

Barrow NJ (2008) The description of sorption curves. Eur J Soil Sci 59:900–910. https://doi.org/10.1111/j.1365-2389.2008.01041.x

Barrow NJ, Shaw TC (1979) Effects of Ionic Strength and Nature of the Cation on Desorption of Phosphate from Soil. Journal of Soil Science 30:53–65. https://doi.org/10.1111/j.1365-2389.1979.tb00964.x

Bartlett R, James B (1980) Studying dried, stored soil samples — Some pitfalls. Soil Science Society of America Journal 44:721. https://doi.org/10.2136/sssaj1980.03615995004400040011x

Battin TJ, Besemer K, Bengtsson MM, et al (2016) The ecology and biogeochemistry of stream biofilms. Nat Rev Microbiol 14:251–263. https://doi.org/10.1038/nrmicro.2016.15

Baulch HM, Futter MN, Jin L, et al (2013) Phosphorus dynamics across intensively monitored subcatchments in the Beaver River. Inland Waters 3:187–206. https://doi.org/10.5268/IW-3.2.530

Bernhardt ES, Heffernan JB, Grimm NB, et al (2018) The metabolic regimes of flowing waters. Limnology and Oceanography 63:S99–S118. https://doi.org/10.1002/lno.10726

Berns AE, Philipp H, Narres H-D, et al (2008) Effect of gamma-sterilization and autoclaving on soil organic matter structure as studied by solid state NMR, UV and fluorescence spectroscopy. European Journal of Soil Science 59:540–550. https://doi.org/10.1111/j.1365-2389.2008.01016.x

Bhadha JH, Daroub SH, Lang TA (2012) Effect of kinetic control, soil: Solution ratio, electrolyte cation, and others, on equilibrium phosphorus concentration. Geoderma 173–174:209–214. https://doi.org/10.1016/j.geoderma.2011.12.027

Bieroza MZ, Heathwaite AL (2015) Seasonal variation in phosphorus concentration–discharge hysteresis inferred from high-frequency in situ monitoring. Journal of Hydrology 524:333–347. https://doi.org/10.1016/j.jhydrol.2015.02.036

Biggs BJ (1996) Patterns in benthic algae of streams. In: Algal Ecology: Freshwater Benthic Ecosystems. pp 31–56

Biggs BJ, Kilroy C (2000) Stream periphyton monitoring manual. NIWA, Christchurch, New Zealand

Biggs BJF (2000) Eutrophication of streams and rivers: dissolved nutrient-chlorophyll relationships for benthic algae. J N Am Benthol Soc 19:17–31. https://doi.org/10.2307/1468279

Biggs BJF, Nikora VI, Snelder TH (2005) Linking scales of flow variability to lotic ecosystem structure and function. River Research and Applications 21:283–298. https://doi.org/10.1002/rra.847

Blakemore LC, Searle PL, Daly BK (1987) Methods for chemical analysis of soils. NZ Soil Bureau, New Zealand

Blankinship JC, Becerra CA, Schaeffer SM, Schimel JP (2014) Separating cellular metabolism from exoenzyme activity in soil organic matter decomposition. Soil Biology and Biochemistry 71:68–75. https://doi.org/10.1016/j.soilbio.2014.01.010

Bluth GJS, Kump LR (1994) Lithologic and climatologic controls of river chemistry. Geochim Cosmochim Acta 58:2341–2359. https://doi.org/10.1016/0016-7037(94)90015-9

Boano F, Harvey JW, Marion A, et al (2014) Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications. Rev Geophys 52:603–679. https://doi.org/10.1002/2012RG000417.Received

Bogert MCV de, Carpenter SR, Cole JJ, Pace ML (2007) Assessing pelagic and benthic metabolism using free water measurements. Limnology and Oceanography: Methods 5:145–155. https://doi.org/10.4319/lom.2007.5.145

Bohrman KJ, Strauss EA (2018) Macrophyte-driven transient storage and phosphorus uptake in a western Wisconsin stream. Hydrological Processes 32:253–263. https://doi.org/10.1002/hyp.11411

Bolan NS, Barrow NJ (1984) Modelling the effect of adsorption of phosphate and other anions on the surface charge of variable charge oxides. Journal of Soil Science 35:273–281. https://doi.org/10.1111/j.1365-2389.1984.tb00282.x

Bolan NS, Barrow NJ, Posner AM (1985) Describing the effect of time on sorption of phosphate by iron and aluminium hydroxides. Journal of Soil Science 36:187–197. https://doi.org/10.1111/j.1365-2389.1985.tb00323.x

Booman GC, Laterra P (2019) Channelizing Streams for Agricultural Drainage Impairs their Nutrient Removal Capacity. Journal of Environmental Quality 48:459–468. https://doi.org/10.2134/jeq2018.07.0264

Borchardt MA (1996) 7 - Nutrients. In: Stevenson RJ, Bothwell ML, Lowe RL (eds) Algal Ecology. Academic Press, San Diego, pp 183–227

Boston HL, Hill WR (1991) Photosynthesis–light relations of stream periphyton communities. Limnology and Oceanography 36:644–656. https://doi.org/10.4319/lo.1991.36.4.0644

Briggs MA, Day-Lewis FD, Zarnetske JP, Harvey JW (2015) A physical explanation for the development of redox microzones in hyporheic flow. Geophys Res Lett 42:4402–4410. https://doi.org/10.1002/2015GL064200

Brookes PC, Powlson DS, Jenkinson DS (1982) Measurement of microbial biomass phosphorus in soil. Soil Biology and Biochemistry 14:319–329

Brooks SC, Brandt CC, Griffiths NA (2017) Estimating uncertainty in ambient and saturation nutrient uptake metrics from nutrient pulse releases in stream ecosystems. Limnology and Oceanography: Methods 15:22–37. https://doi.org/10.1002/lom3.10139

Brown LJ (2001) Canterbury. In: Groundwaters of New Zealand. New Zealand Hydrological Society Inc., Wellington, New Zealand, pp 441–459

Buessecker S, Tylor K, Nye J, et al (2019) Effects of sterilization techniques on chemodenitrification and N_2O production in tropical peat soil microcosms. Biogeosciences 16:4601–4612. https://doi.org/10.5194/bg-16-4601-2019

Burkitt LL, Moody PW, Gourley CJP, Hannah MC (2002) A simple phosphorus buffering index for Australian soils. Soil Res 40:497–513. https://doi.org/10.1071/sr01050

Burt TP, Pinay G (2005) Linking hydrology and biogeochemistry in complex landscapes. Progress in Physical Geography: Earth and Environment 29:297–316. https://doi.org/10.1191/0309133305pp450ra

Cade-Menun BJ (2005) Characterizing phosphorus in environmental and agricultural samples by 31P nuclear magnetic resonance spectroscopy. Talanta 66:359–371. https://doi.org/10.1016/j.talanta.2004.12.024

Campisano R, Hall K, Griggs J, et al (2017) Selected Analytical Methods for Environmental Remediation and Recovery (SAM) 2017. U.S. Environmental Protection Agency, Washington, D.C.

Carter MR, Gregorich EG (2007) Soil Sampling and Methods of Analysis. CRC Press

Casillas-Ituarte NN, Sawyer AH, Danner KM, et al (2020) Internal Phosphorus Storage in Two Headwater Agricultural Streams in the Lake Erie Basin. Environ Sci Technol 54:176–183. https://doi.org/10.1021/acs.est.9b04232

Church M (2002) Geomorphic thresholds in riverine landscapes. Freshwater Biol 47:541–557. https://doi.org/10.1046/j.1365-2427.2002.00919.x

Cleveland CC, Liptzin D (2007) C:N:P stoichiometry in soil: is there a “Redfield ratio” for the microbial biomass? Biogeochemistry 85:235–252. https://doi.org/10.1007/s10533-007-9132-0

Coelho JP, Flindt MR, Jensen HS, et al (2004) Phosphorus speciation and availability in intertidal sediments of a temperate estuary: relation to eutrophication and annual P-fluxes. Estuarine Coastal Shelf Sci 61:583–590. https://doi.org/10.1016/j.ecss.2004.07.001

Cohen MJ, Kurz MJ, Heffernan JB, et al (2013) Diel phosphorus variation and the stoichiometry of ecosystem metabolism in a large spring-fed river. Ecol Monogr 83:155–176. https://doi.org/10.1890/12-1497.1

Condron LM, Newman S (2011) Revisiting the fundamentals of phosphorus fractionation of sediments and soils. J Soils Sediments 11:830–840. https://doi.org/10.1007/s11368-011-0363-2

Corman JR, Moody EK, Elser JJ (2016) Calcium carbonate deposition drives nutrient cycling in a calcareous headwater stream. Ecological Monographs 86:448–461. https://doi.org/10.1002/ecm.1229

Corman JR, Moody EK, Elser JJ (2015) Stoichiometric impact of calcium carbonate deposition on nitrogen and phosphorus supplies in three montane streams. Biogeochemistry 126:285–300. https://doi.org/10.1007/s10533-015-0156-6

Côté IM, Curtis PS, Rothstein HR, Stewart GB (2013) Gathering data: searching literature and selection criteria. In: Handbook of Meta-analysis in Ecology and Evolution. Princeton University Press, Princeton, New Jersey, pp 37–51

Covino T (2017) Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. Geomorphology 277:133–144. https://doi.org/10.1016/j.geomorph.2016.09.030

Covino TP, McGlynn BL, McNamara RA (2010) Tracer Additions for Spiraling Curve Characterization (TASCC): Quantifying stream nutrient uptake kinetics from ambient to saturation. Limnol Oceanogr Methods 8:484–498. https://doi.org/10.4319/lom.2010.8.484

Crane-Droesch A, Abiven S, Jeffery S, Torn MS (2013) Heterogeneous global crop yield response to biochar: a meta-regression analysis. Environ Res Lett 8:044049. https://doi.org/10.1088/1748-9326/8/4/044049

Crockford L, Jordan P, Melland AR, Taylor D (2015) Storm-triggered, increased supply of sediment-derived phosphorus to the epilimnion in a small freshwater lake. Inland Waters 5:15–26. https://doi.org/10.5268/IW-5.1.738

Cross WF, Benstead JP, Frost PC, Thomas SA (2005) Ecological stoichiometry in freshwater benthic systems: recent progress and perspectives. Freshwater Biology 50:1895–1912. https://doi.org/10.1111/j.1365-2427.2005.01458.x

Dahlgren RA, Saigusa M, Ugolini FC (2004) The Nature, Properties and Management of Volcanic Soils. In: Advances in Agronomy. Academic Press, pp 113–182

Danen-Louwerse H, Lijklema L, Coenraats M (1993) Iron content of sediment and phosphate adsorption properties. Hydrobiol 253:311–317

D'Angelo E, Crutchfield J, Vandiviere M (2001) Rapid, sensitive, microscale determination of phosphate in water and soil. J Environ Qual 30:2206–2209. https://doi.org/10.2134/jeq2001.2206

Demars BOL (2019) Hydrological pulses and burning of dissolved organic carbon by stream respiration. Limnol Oceanogr 64:406–421. https://doi.org/10.1002/lno.11048

Demars BOL (2008) Whole-stream phosphorus cycling: Testing methods to assess the effect of saturation of sorption capacity on nutrient uptake length measurements. Water Res 42:2507–2516. https://doi.org/10.1016/j.watres.2008.02.010

Diaz OA, Reddy KR, Moore PA (1994) Solubility of inorganic phosphorus in stream water as influenced by pH and calcium concentration. Water Res 28:1755–1763. https://doi.org/10.1016/0043-1354(94)90248-8

Dick WA, Tabatabai MA (1977) Determination of orthophosphate in aqueous solutions containing labile organic and inorganic phosphorus compounds. Journal of Environment Quality 6:82–85. https://doi.org/10.2134/jeq1977.00472425000600010018x

Dieter D, Herzog C, Hupfer M (2015) Effects of drying on phosphorus uptake in re-flooded lake sediments. Environmental Science and Pollution Research 22:17065–17081. https://doi.org/10.1007/s11356-015-4904-x

Dietrich WE (1989) Sediment supply and the development of the coarse surface layer in gravel-bedded rivers. Nature 340:215–217

do Nascimento CAC, Pagliari PH, Schmitt D, et al (2016) Phosphorus Concentrations in Sequentially Fractionated Soil Samples as Affected by Digestion Methods. Scientific Reports 5:17967. https://doi.org/10.1038/srep17967

Dodds WK (2007) Trophic state, eutrophication and nutrient criteria in streams. Trends in Ecology & Evolution 22:669–676. https://doi.org/10.1016/j.tree.2007.07.010

Dodds WK (2003) The Role of Periphyton in Phosphorus Retention in Shallow Freshwater Aquatic Systems. Journal of Phycology 39:840–849. https://doi.org/10.1046/j.1529-8817.2003.02081.x

Dodds WK (2006) Eutrophication and trophic state in rivers and streams. Limnol Oceanogr 51:671–680. https://doi.org/10.4319/lo.2006.51.1_part_2.0671

Dodds WK, Bouska WW, Eitzmann JL, et al (2009) Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages. Environ Sci Technol 43:12–19. https://doi.org/10.1021/es801217q

Dodds WK, Smith VH (2016) Nitrogen, phosphorus, and eutrophication in streams. Inland Waters 6:155–164. https://doi.org/10.5268/IW-6.2.909

Dodds WK, Smith VH, Lohman K (2002) Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Can J Fish Aquat Sci 59:865–874. https://doi.org/10.1139/f02-063

Drohan PJ, Bechmann M, Buda A, et al (2019) A Global Perspective on Phosphorus Management Decision Support in Agriculture: Lessons Learned and Future Directions. Journal of Environmental Quality 48:1218–1233. https://doi.org/10.2134/jeq2019.03.0107

Drummond JD, Covino TP, Aubeneau AF, et al (2012) Effects of solute breakthrough curve tail truncation on residence time estimates: A synthesis of solute tracer injection studies. J Geophys Res Biogeosci 117:. https://doi.org/10.1029/2012JG002019

Dupas R, Abbott BW, Minaudo C, Fovet O (2019) Distribution of Landscape Units Within Catchments Influences Nutrient Export Dynamics. Front Environ Sci 7:. https://doi.org/10.3389/fenvs.2019.00043

Dupas R, Minaudo C, Gruau G, et al (2018a) Multidecadal trajectory of riverine nitrogen and phosphorus dynamics in rural catchments. Water Resour Res 54:5327–5340. https://doi.org/10.1029/2018WR022905

Dupas R, Musolff A, Jawitz JW, et al (2017) Carbon and nutrient export regimes from headwater catchments to downstream reaches. Biogeosciences 14:4391–4407. https://doi.org/10.5194/bg-14-4391-2017

Dupas R, Tittel J, Jordan P, et al (2018b) Non-domestic phosphorus release in rivers during low-flow: Mechanisms and implications for sources identification. J Hydrol 560:141–149. https://doi.org/10.1016/j.jhydrol.2018.03.023

Dzombak DA, Morel FMM (1987) Adsorption of Inorganic Pollutants in Aquatic Systems. Journal of Hydraulic Engineering 113:430–475. https://doi.org/10.1061/(ASCE)0733-9429(1987)113:4(430)

Ekka SA, Haggard BE, Matlock MD, Chaubey I (2006) Dissolved phosphorus concentrations and sediment interactions in effluent-dominated Ozark streams. Ecol Eng 26:375–391. https://doi.org/10.1016/j.ecoleng.2006.01.002

Elser JJ, Bracken MES, Cleland EE, et al (2007) Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecology Letters 10:1135–1142. https://doi.org/10.1111/j.1461-0248.2007.01113.x

Elwood JW, Mulholland PJ, Newbold JD (1988) Microbial activity and phosphorus uptake on decomposing leaf detritus in a heterotrophic stream. SIL Proceedings, 1922-2010 23:1198–1208. https://doi.org/10.1080/03680770.1987.11899793

Emelko MB, Stone M, Silins U, et al (2016) Sediment-phosphorus dynamics can shift aquatic ecology and cause downstream legacy effects after wildfire in large river systems. Global Change Biol 22:1168–1184. https://doi.org/10.1111/gcb.13073

Eno CF, Popenoe H (1964) Gamma Radiation Compared with Steam and Methyl Bromide as a Soil Sterilizing Agent. Soil Science Society of America Journal 28:533–535. https://doi.org/10.2136/sssaj1964.03615995002800040024x

Ensign SH, Doyle MW (2006) Nutrient spiraling in streams and river networks. Journal of Geophysical Research: Biogeosciences 111:. https://doi.org/10.1029/2005JG000114

Eshel G, Levy GJ, Mingelgrin U, Singer MJ (2004) Critical evaluation of the use of laser diffraction for particle-size distribution analysis. Soil Sci Soc Am J 68:736–743. https://doi.org/10.2136/sssaj2004.7360

Evans-White MA, Dodds WK, Huggins DG, Baker DS (2009) Thresholds in macroinvertebrate biodiversity and stoichiometry across water-quality gradients in Central Plains (USA) streams. Journal of the North American Benthological Society 28:855–868. https://doi.org/10.1899/08-113.1

Ezzati G, Fenton O, Healy MG, et al (2020) Impact of P inputs on source-sink P dynamics of sediment along an agricultural ditch network. Journal of Environmental Management 257:109988. https://doi.org/10.1016/j.jenvman.2019.109988

Falco ND, Boano F, Arnon S (2016) Biodegradation of labile dissolved organic carbon under losing and gaining streamflow conditions simulated in a laboratory flume. Limnology and Oceanography 61:1839–1852. https://doi.org/10.1002/lno.10344

Fanelli RM, Blomquist JD, Hirsch RM (2019) Point sources and agricultural practices control spatial-temporal patterns of orthophosphate in tributaries to Chesapeake Bay. Science of The Total Environment 652:422–433. https://doi.org/10.1016/j.scitotenv.2018.10.062

Fanta SE, Hill WR, Smith TB, Roberts BJ (2010) Applying the light : nutrient hypothesis to stream periphyton. Freshwater Biology 55:931–940. https://doi.org/10.1111/j.1365-2427.2009.02309.x

Femeena PV, Chaubey I, Aubeneau A, et al (2019) Simple regression models can act as calibration-substitute to approximate transient storage parameters in streams. Advances in Water Resources 123:201–209. https://doi.org/10.1016/j.advwatres.2018.11.010

Findlay S (2016) Chapter 3 - Stream Microbial Ecology in a Changing Environment. In: Jones JB, Stanley EH (eds) Stream Ecosystems in a Changing Environment. Academic Press, Boston, pp 135–150

Findlay S (1995) Importance of surface-subsurface exchange in stream ecosystems: The hyporheic zone. Limnol Oceanogr 40:159–164. https://doi.org/10.4319/lo.1995.40.1.0159

Findlay S, Sinsabaugh RL (2006) Large-Scale Variation in Subsurface Stream Biofilms: A Cross-Regional Comparison of Metabolic Function and Community Similarity. Microb Ecol 52:491–500. https://doi.org/10.1007/s00248-006-9095-z

Findlay SEG, Sinsabaugh RL, Sobczak WV, Hoostal M (2003) Metabolic and structural response of hyporheic microbial communities to variations in supply of dissolved organic matter. Limnology and Oceanography 48:1608–1617. https://doi.org/10.4319/lo.2003.48.4.1608

Finlay JC (2011) Stream size and human influences on ecosystem production in river networks. Ecosphere 2:. https://doi.org/10.1890/ES11-00071.1

Finlay JC, Hood JM, Limm MP, et al (2011) Light-mediated thresholds in stream-water nutrient composition in a river network. Ecology 92:140–150. https://doi.org/10.1890/09-2243.1

Fox LE (1988) The solubility of colloidal ferric hydroxide and its relevance to iron concentrations in river water. Geochim Cosmochim Acta 52:771–777. https://doi.org/10.1016/0016-7037(88)90337-7

Frei RJ, Abbott BW, Dupas R, et al (2020) Predicting Nutrient Incontinence in the Anthropocene at Watershed Scales. Front Environ Sci 7:. https://doi.org/10.3389/fenvs.2019.00200

Froelich PN (1988) Kinetic control of dissolved phosphate in natural rivers and estuaries: A primer on the phosphate buffer mechanism. Limnol Oceanogr 33:649–668

Gainswin BE, House WA, Leadbeater BSC, et al (2006) The effects of sediment size fraction and associated algal biofilms on the kinetics of phosphorus release. Sci Total Environ 360:142–157. https://doi.org/10.1016/j.scitotenv.2005.08.034

Gebremikael MT, De Waele J, Buchan D, et al (2015) The effect of varying gamma irradiation doses and soil moisture content on nematodes, the microbial communities and mineral nitrogen. Applied Soil Ecology 92:1–13. https://doi.org/10.1016/j.apsoil.2015.03.003

Gérard F (2016) Clay minerals, iron/aluminum oxides, and their contribution to phosphate sorption in soils — A myth revisited. Geoderma 262:213–226. https://doi.org/10.1016/j.geoderma.2015.08.036

Goedkoop W, Pettersson K (2000) Seasonal changes in sediment phosphorus forms in relation to sedimentation and benthic bacterial biomass in Lake Erken. Hydrobiologia 431:41–50. https://doi.org/10.1023/A:1004050204587

Goldberg S, Sposito G (1985) On the mechanism of specific phosphate adsorption by hydroxylated mineral surfaces: A review. Commun Soil Sci Plant Anal 16:801–821. https://doi.org/10.1080/00103628509367646

Goldberg S, Sposito G (1984a) A Chemical Model of Phosphate Adsorption by Soils: II. Noncalcareous Soils 1. Soil Sci Soc Am J 48:779–783. https://doi.org/10.2136/sssaj1984.03615995004800040016x

Goldberg S, Sposito G (1984b) A chemical model of phosphate adsorption by soils: I. Reference oxide minerals. Soil Sci Soc Am J 48:772–778. https://doi.org/10.2136/sssaj1984.03615995004800040015x

Golterman HL (1996) Fractionation of sediment phosphate with chelating compounds: a simplification, and comparison with other methods. Hydrobiol 335:87–95. https://doi.org/10.1007/BF00013687

Golterman HL (2004) The chemistry of phosphate and nitrogen compounds in sediments. Kluwer Academic Publishers, Dordrecht, The Netherlands

Golterman HL (1988) The calcium- and iron bound phosphate phase diagram. Hydrobiol 159:149–151. https://doi.org/10.1007/BF00014722

Golterman HL (2002) Editorial. Hydrobiologia 472:3–4. https://doi.org/10.1023/A:1016319027731

Gottselig N, Amelung W, Kirchner JW, et al (2017) Elemental composition of natural nanoparticles and fine colloids in European forest stream waters and their role as phosphorus carriers. Global Biogeochem Cycles 31:1592–1607. https://doi.org/10.1002/2017GB005657

Gran KB, Czuba JA (2017) Sediment pulse evolution and the role of network structure. Geomorphology 277:17–30. https://doi.org/10.1016/j.geomorph.2015.12.015

Griffin BA, Jurinak JJ (1973) Estimation Of Activity Coefficients From The Electrical Conductivity Of Natural Aquatic Systems And Soil Extracts. Soil Science 116:26–30

Griffiths NA, Johnson LT (2018) Influence of dual nitrogen and phosphorus additions on nutrient uptake and saturation kinetics in a forested headwater stream. Freshwater Science 37:810–825. https://doi.org/10.1086/700700

Guan XH, Shang C, Chen GH (2006) Competitive adsorption of organic matter with phosphate on aluminum hydroxide. Journal of Colloid and Interface Science 296:51–58. https://doi.org/10.1016/j.jcis.2005.08.050

Gurevitch J, Koricheva J, Nakagawa S, Stewart G (2018) Meta-analysis and the science of research synthesis. Nature 555:175–182. https://doi.org/10.1038/nature25753

Gustafsson JP, Mwamila LB, Kergoat K (2012) The pH dependence of phosphate sorption and desorption in Swedish agricultural soils. Geoderma 189–190:304–311. https://doi.org/10.1016/j.geoderma.2012.05.014

Haggard BE (2010) Phosphorus Concentrations, Loads, and Sources within the Illinois River Drainage Area, Northwest Arkansas, 1997–2008. Journal of Environmental Quality 39:2113–2120. https://doi.org/10.2134/jeq2010.0049

Haggard BE, Ekka SA, Matlock MD, Chaubey I (2004) Phosphate equilibrium between stream sediments and water: Potential effect of chemical amendments. Transactions of the American Society of Agricultural Engineers 47:1113–1118

Haggard BE, Sharpley AN (2007) Phosphorus transport in streams: Processes and modeling considerations. In: Modeling phosphorus in the environment. CRC Press, Boca Raton, FL

Haggard BE, Smith DR, Brye KR (2007) Variations in stream water and sediment phosphorus among select Ozark catchments. J Environ Qual 36:1725–1734. https://doi.org/10.2134/jeq2006.0517

Haggard BE, Stanley EH, Hyler R (1999) Sediment-phosphorus relationships in three northcentral Oklahoma streams. Transactions of the ASAE 42:1709–1714

Haggard BE, Stanley EH, Storm DE (2005) Nutrient retention in a point-source-enriched stream. J N Am Benthol Soc 24:29–47. https://doi.org/10.1899/0887-3593(2005)024<0029:NRIAPS>2.0.CO;2

Haggerty R, Martí E, Argerich A, et al (2009) Resazurin as a “smart” tracer for quantifying metabolically active transient storage in stream ecosystems. Journal of Geophysical Research: Biogeosciences 114:. https://doi.org/10.1029/2008JG000942

Hall RO, Baker MA, Rosi-Marshall EJ, et al (2013) Solute-specific scaling of inorganic nitrogen and phosphorus uptake in streams. Biogeosciences 10:7323–7331. https://doi.org/10.5194/bg-10-7323-2013

Hall RO, Bernhardt ES, Likens GE (2002) Relating nutrient uptake with transient storage in forested mountain streams. Limnol Oceanogr 47:255–265. https://doi.org/10.4319/lo.2002.47.1.0255

Hall RO, Hotchkiss ER (2017) Chapter 34 - Stream Metabolism. In: Lamberti GA, Hauer FR (eds) Methods in Stream Ecology (Third Edition). Academic Press, pp 219–233

Hall RO, Tank JL (2005) Correcting whole-stream estimates of metabolism for groundwater input. Limnology and Oceanography: Methods 3:222–229. https://doi.org/10.4319/lom.2005.3.222

Hall RO, Tank JL, Baker MA, et al (2016) Metabolism, Gas Exchange, and Carbon Spiraling in Rivers. Ecosystems 19:73–86. https://doi.org/10.1007/s10021-015-9918-1

Hamilton SK (2012) Biogeochemical time lags may delay responses of streams to ecological restoration. Freshwater Biol 57:43–57. https://doi.org/10.1111/j.1365-2427.2011.02685.x

Hanrahan BR, Tank JL, Aubeneau AF, Bolster D (2018) Substrate-specific biofilms control nutrient uptake in experimental streams. Freshwater Science 37:456–471. https://doi.org/10.1086/699004

Harding EF (1986) Modelling: the classical approach. Journal of the Royal Statistical Society 35:115–134

Hartwig M, Borchardt D (2015) Alteration of key hyporheic functions through biological and physical clogging along a nutrient and fine-sediment gradient. Ecohydrology 8:961–975. https://doi.org/10.1002/eco.1571

Harvey JW (2016) Hydrologic exchange flows and their ecological consequences in river corridors. In: Stream Ecosystems in a Changing Environment. Academic Press, London, UK, pp 1–84

Harvey JW, Böhlke JK, Voytek MA, et al (2018) Hyporheic zone denitrification: Controls on effective reaction depth and contribution to whole-stream mass balance. Water Resources Research 6298–6316. https://doi.org/10.1002/wrcr.20492@10.1002/(ISSN)1944-7973.AQUIDER1

Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, 2nd edn. Springer-Verlag, New York

Haygarth PM, Warwick MS, House WA (1997) Size distribution of colloidal molybdate reactive phosphorus in river waters and soil solution. Water Res 31:439–448. https://doi.org/10.1016/S0043-1354(96)00270-9

He Z, Honeycutt CW (2005) A modified molybdenum blue method for orthophosphate determination suitable for investigating enzymatic hydrolysis of organic phosphates. Commun Soil Sci Plant Anal 36:1373–1383. https://doi.org/10.1081/CSS-200056954

Helsel DR (2005) More Than Obvious: Better Methods for Interpreting Nondetect Data. Environ Sci Technol 39:419A-423A. https://doi.org/10.1021/es053368a

Henningsen A, Toomet O (2011) maxLik: A package for maximum likelihood estimation in R. Comput Stat 26:443–458. https://doi.org/10.1007/s00180-010-0217-1

Hensley RT, Cohen MJ (2016) On the emergence of diel solute signals in flowing waters. Water Resources Research 52:759–772. https://doi.org/10.1002/2015WR017895

Henson E, Lasater A, Haggard BE (2019) Reducing Dissolved Phosphorus in Stream Water May Not Influence Estimation of Sediment Equilibrium Phosphorus Concentrations. Agrosystems, Geosciences & Environment 2:. https://doi.org/10.2134/age2019.05.0037

Herndon EM, Kinsman-Costello L, Duroe KA, et al (2019) Iron (oxyhydr)oxides serve as phosphate traps in tundra and boreal peat soils. J Geophys Res Biogeosci 124:. https://doi.org/10.1029/2018JG004776

Hesterberg D (2010) Chapter 11 - Macroscale Chemical Properties and X-Ray Absorption Spectroscopy of Soil Phosphorus. In: Singh B, Gräfe M (eds) Developments in Soil Science. Elsevier, pp 313–356

Hewitt AE (2010) New Zealand Soil Classification (NZSC), 3rd edn. Manaaki Whenua Press, Landcare Research, Lincoln, New Zealand

Hieltjes AHM, Lijklema L (1980) Fractionation of Inorganic Phosphates in Calcareous Sediments. Journal of Environmental Quality 9:405–407. https://doi.org/10.2134/jeq1980.00472425000900030015x

Hill BH, Elonen CM, Jicha TM, et al (2010) Sediment microbial enzyme activity as an indicator of nutrient limitation in the great rivers of the Upper Mississippi River basin. Biogeochemistry 97:195–209. https://doi.org/10.1007/s10533-009-9366-0

Hill BH, Elonen CM, Seifert LR, et al (2012) Microbial enzyme stoichiometry and nutrient limitation in US streams and rivers. Ecological Indicators 18:540–551. https://doi.org/10.1016/j.ecolind.2012.01.007

Hill BH, Herlihy AT, Kaufmann PR (2002) Benthic microbial respiration in Appalachian Mountain, Piedmont, and Coastal Plains streams of the eastern U.S.A. Freshwater Biology 47:185–194. https://doi.org/10.1046/j.1365-2427.2002.00791.x

Hill WR, Fanta SE, Roberts BJ (2009) Quantifying phosphorus and light effects in stream algae. Limnology and Oceanography 54:368–380. https://doi.org/10.4319/lo.2009.54.1.0368

Hirabayashi Y, Mahendran R, Koirala S, et al (2013) Global flood risk under climate change. Nature Clim Change 3:816–821. https://doi.org/10.1038/nclimate1911

Hjorth T (2004) Effects of freeze-drying on partitioning patterns of major elements and trace metals in lake sediments. Analytica Chimica Acta 526:95–102. https://doi.org/10.1016/j.aca.2004.08.007

Hoffman AR, Armstrong DE, Lathrop RC, Penn MR (2009) Characteristics and influence of phosphorus accumulated in the bed sediments of a stream located in an agricultural watershed. Aquat Geochem 15:371–389. https://doi.org/10.1007/s10498-008-9043-2

Hollander M, Wolfe DA, Chicken E (2013a) The one-sample location problem. In: Nonparametric Statistical Methods, 3rd edn. John Wiley & Sons, pp 39–114

Hollander M, Wolfe DA, Chicken E (2013b) The one-way layout. In: Nonparametric Statistical Methods, 3rd edn. John Wiley & Sons, USA, pp 202–288

Holtgrieve GW, Schindler DE, Branch TA, A'mar ZT (2010) Simultaneous quantification of aquatic ecosystem metabolism and reaeration using a Bayesian statistical model of oxygen dynamics. Limnology and Oceanography 55:1047–1063. https://doi.org/10.4319/lo.2010.55.3.1047

Hongthanat N, Kovar JL, Thompson ML, et al (2016) Phosphorus source—sink relationships of stream sediments in the Rathbun Lake watershed in southern Iowa, USA. Environmental Monitoring and Assessment 188:. https://doi.org/10.1007/s10661-016-5437-6

Horn DJV, Sinsabaugh RL, Takacs-Vesbach CD, et al (2011) Response of heterotrophic stream biofilm -communities to a gradient of resources. Aquatic Microbial Ecology 64:149–161. https://doi.org/10.3354/ame01515

House WA (2003) Geochemical cycling of phosphorus in rivers. Appl Geochem 18:739–748. https://doi.org/10.1016/S0883-2927(02)00158-0

House WA (1999) The physico-chemical conditions for the precipitation of phosphate with calcium. Environ Technol 20:727–733. https://doi.org/10.1080/09593332008616867

House WA, Denison FH (2002) Exchange of inorganic phosphate between river waters and bed-sediments. Environ Sci Technol 36:4295–4301. https://doi.org/10.1021/es020039z

House WA, Denison FH (1997) Nutrient dynamics in a lowland stream impacted by sewage effluent: Great Ouse, England. Sci Total Environ 205:25–49. https://doi.org/10.1016/S0048-9697(97)00086-7

House WA, Denison FH (2000) Factors influencing the measurement of equilibrium phosphate concentrations in river sediments. Water Res 34:1187–1200. https://doi.org/10.1016/S0043-1354(99)00249-3

House WA, Denison FH (1998) Phosphorus dynamics in a lowland river. Water Research 32:1819–1830. https://doi.org/10.1016/S0043-1354(97)00407-7

House WA, Warwick MS (1999) Interactions of phosphorus with sediments in the River Swale, Yorkshire, UK. Hydrol Processes 13:1103–1115. https://doi.org/10.1002/(SICI)1099-1085(199905)13:7<1103::AID-HYP792>3.0.CO;2-6

Huang L, Fang H, He G, Chen M (2016) Phosphorus adsorption on natural sediments with different pH incorporating surface morphology characterization. Environ Sci Technol 23:18883–18891. https://doi.org/10.1007/s11356-016-7093-3

Hyacinthe C, Bonneville S, Van Cappellen P (2006) Reactive iron(III) in sediments: Chemical versus microbial extractions. Geochim Cosmochim Acta 70:4166–4180. https://doi.org/10.1016/j.gca.2006.05.018

Jackson-Blake LA, Dunn SM, Helliwell RC, et al (2015) How well can we model stream phosphorus concentrations in agricultural catchments? Environmental Modelling & Software 64:31–46. https://doi.org/10.1016/j.envsoft.2014.11.002

Jaisi DP, Blake RE (2014) Chapter One - Advances in Using Oxygen Isotope Ratios of Phosphate to Understand Phosphorus Cycling in the Environment. In: Sparks DL (ed) Advances in Agronomy. Academic Press, pp 1–53

Jaiswal D, Pandey J (2019) Investigations on peculiarities of land-water interface and its use as a stable testbed for accurately predicting changes in ecosystem responses to human perturbations: A sub-watershed scale study with the Ganga River. Journal of Environmental Management 238:178–193. https://doi.org/10.1016/j.jenvman.2019.02.126

Jan J, Borovec J, Kopáček J, Hejzlar J (2015) Assessment of phosphorus associated with Fe and Al (hydr)oxides in sediments and soils. J Soils Sediments 15:1620–1629. https://doi.org/10.1007/s11368-015-1119-1

Jan J, Borovec J, Kopáček J, Hejzlar J (2013) What do results of common sequential fractionation and single-step extractions tell us about P binding with Fe and Al compounds in non-calcareous sediments? Water Res 47:547–557. https://doi.org/10.1016/j.watres.2012.10.053

Jankowski K, Schindler DE, Lisi PJ (2014) Temperature sensitivity of community respiration rates in streams is associated with watershed geomorphic features. Ecology 95:2707–2714. https://doi.org/10.1890/14-0608.1

Jansson M, Berggren M, Laudon H, Jonsson A (2012) Bioavailable phosphorus in humic headwater streams in boreal Sweden. Limnology and Oceanography 57:1161–1170. https://doi.org/10.4319/lo.2012.57.4.1161

Jarvie HP, Jürgens MD, Williams RJ, et al (2005) Role of river bed sediments as sources and sinks of phosphorus across two major eutrophic UK river basins: the Hampshire Avon and Herefordshire Wye. J Hydrol 304:51–74. https://doi.org/10.1016/j.jhydrol.2004.10.002

Jarvie HP, Neal C, Jürgens MD, et al (2006a) Within-river nutrient processing in Chalk streams: The Pang and Lambourn, UK. J Hydrol 330:101–125. https://doi.org/10.1016/j.jhydrol.2006.04.014

Jarvie HP, Neal C, Withers PJA (2006b) Sewage-effluent phosphorus: A greater risk to river eutrophication than agricultural phosphorus? Sci Total Environ 360:246–253. https://doi.org/10.1016/j.scitotenv.2005.08.038

Jarvie HP, Sharpley AN, Flaten D, Kleinman PJA (2019) Phosphorus mirabilis: Illuminating the Past and Future of Phosphorus Stewardship. Journal of Environmental Quality 48:1127–1132. https://doi.org/10.2134/jeq2019.07.0266

Jarvie HP, Sharpley AN, Kresse T, et al (2018) Coupling High-Frequency Stream Metabolism and Nutrient Monitoring to Explore Biogeochemical Controls on Downstream Nitrate Delivery. Environ Sci Technol 52:13708–13717. https://doi.org/10.1021/acs.est.8b03074

Jarvie HP, Sharpley AN, Scott JT, et al (2012) Within-river phosphorus retention: accounting for a missing piece in the watershed phosphorus puzzle. Environ Sci Technol 46:13284–13292. https://doi.org/10.1021/es303562y

Jarvie HP, Sharpley AN, Spears B, et al (2013a) Water Quality Remediation Faces Unprecedented Challenges from "Legacy Phosphorus." Environ Sci Technol 47:8997–8998. https://doi.org/10.1021/es403160a

Jarvie HP, Sharpley AN, Withers PJA, et al (2013b) Phosphorus Mitigation to Control River Eutrophication: Murky Waters, Inconvenient Truths, and "Postnormal" Science. Journal of Environmental Quality 42:295–304. https://doi.org/10.2134/jeq2012.0085

Jassby AD, Platt T (1976) Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. Limnology and Oceanography 21:540–547. https://doi.org/10.4319/lo.1976.21.4.0540

Jenkinson DS, Brookes PC, Powlson DS (2004) Measuring soil microbial biomass. Soil Biology and Biochemistry 36:5–7. https://doi.org/10.1016/j.soilbio.2003.10.002

Jensen HS, Kristensen P, Jeppesen E, Skytthe A (1992) Iron:phosphorus ratio in surface sediment as an indicator of phosphate release from aerobic sediments in shallow lakes. In: Hart BT, Sly PG (eds) Sediment/Water Interactions. Springer Netherlands, The Netherlands, pp 731–743

Jensen HS, McGlathery KJ, Marino R, Howarth RW (1998) Forms and availability of sediment phosphorus in carbonate sand of Bermuda seagrass beds. Limnology and Oceanography 43:799–810. https://doi.org/10.4319/lo.1998.43.5.0799

Jensen HS, Thamdrup B (1993) Iron-bound phosphorus in marine sediments as measured by bicarbonate-dithionite extraction. In: Boers PCM, Cappenberg ThE, van Raaphorst W (eds) Proceedings of the Third International Workshop on Phosphorus in Sediments. Springer Netherlands, pp 47–59

Johnson LT, Tank JL (2009) Diurnal variations in dissolved organic matter and ammonium uptake in six open-canopy streams. Journal of the North American Benthological Society 28:694–708. https://doi.org/10.1899/08-107.1

Jolani S, Debray TPA, Koffijberg H, et al (2015) Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. Statistics in Medicine 34:1841–1863. https://doi.org/10.1002/sim.6451

Julian JP, Podolak CJP, Meitzen KM, et al (2016) Chapter 2 - Shaping the Physical Template: Biological, Hydrological, and Geomorphic Connections in Stream Channels. In: Jones JB, Stanley EH (eds) Stream Ecosystems in a Changing Environment. Academic Press, Boston, pp 85–133

Kamerlin SCL, Sharma PK, Prasad RB, Warshel A (2013) Why nature really chose phosphate. Quarterly Reviews of Biophysics 46:1–132. https://doi.org/10.1017/S0033583512000157

Kaplan LA, Cory RM (2016) Chapter 6 - Dissolved Organic Matter in Stream Ecosystems: Forms, Functions, and Fluxes of Watershed Tea. In: Jones JB, Stanley EH (eds) Stream Ecosystems in a Changing Environment. Academic Press, Boston, pp 241–320

Kerr JG, Burford M, Olley J, Udy J (2010) The effects of drying on phosphorus sorption and speciation in subtropical river sediments. Mar Freshwater Res 61:928–935. https://doi.org/10.1071/MF09124

Khare N, Hesterberg D, Martin JD (2005) XANES Investigation of Phosphate Sorption in Single and Binary Systems of Iron and Aluminum Oxide Minerals. Environ Sci Technol 39:2152–2160. https://doi.org/10.1021/es049237b

Klotz RL (1988) Sediment control of soluble reactive phosphorus in Hoxie Gorge Creek, New York. Canadian Journal of Fisheries and Aquatic Sciences 45:2026–2034

Klotz RL (1985) Factors controlling phosphorus limitation in stream sediments. Limnol Oceanogr 30:543–553. https://doi.org/10.4319/lo.1985.30.3.0543

Klotz RL (1991) Temporal relation between soluble reactive phosphorus and factors in stream water and sediments in Hoxie Gorge Creek, New York. Canadian Journal of Fisheries and Aquatic Sciences 48:84–90. https://doi.org/10.1139/f91-012

Klotz RL, Linn SA (2001) Influence of Factors Associated with Water Level Drawdown on Phosphorus Release from Sediments. Lake and Reservoir Management 17:48–54. https://doi.org/10.1080/07438140109353972

Konietschke F, Placzek M, Schaarschmidt F, Hothorn LA (2015) nparcomp: An R software package for nonparametric multiple comparisons and simultaneous confidence intervals. J Stat Software 64:1–17. https://doi.org/10.18637/jss.v064.i09

Koricheva J, Gurevitch J (2014) Uses and misuses of meta-analysis in plant ecology. Journal of Ecology 102:828–844. https://doi.org/10.1111/1365-2745.12224

Kouno K, Tuchiya Y, Ando T (1995) Measurement of soil microbial biomass phosphorus by an anion exchange membrane method. Soil Biology and Biochemistry 27:1353–1357. https://doi.org/10.1016/0038-0717(95)00057-L

Kreiling RM, Thoms MC, Bartsch LA, et al (2019) Complex Response of Sediment Phosphorus to Land Use and Management Within a River Network. Journal of Geophysical Research: Biogeosciences 124:1764–1780. https://doi.org/10.1029/2019JG005171

Kreiling RM, Thoms MC, Bartsch LA, et al (2020) Land Use Effects on Sediment Nutrient Processes in a Heavily Modified Watershed Using Structural Equation Models. Water Resources Research 56:e2019WR026655. https://doi.org/10.1029/2019WR026655

Kröger R, Moore MT (2011) Phosphorus dynamics within agricultural drainage ditches in the lower Mississippi Alluvial Valley. Ecological Engineering 37:1905–1909. https://doi.org/10.1016/j.ecoleng.2011.06.042

Kronvang B, Hoffmann CC, Dröge R (2009) Sediment deposition and net phosphorus retention in a hydraulically restored lowland river floodplain in Denmark: combining field and laboratory experiments. Mar Freshwater Res 60:638–646. https://doi.org/10.1071/MF08066

Krumina L, Kenney JPL, Loring JS, Persson P (2016) Desorption mechanisms of phosphate from ferrihydrite and goethite surfaces. Chem Geol 427:54–64. https://doi.org/10.1016/j.chemgeo.2016.02.016

Kusmer AS, Goyette J-O, MacDonald GK, et al (2019) Watershed Buffering of Legacy Phosphorus Pressure at a Regional Scale: A Comparison Across Space and Time. Ecosystems 22:91–109. https://doi.org/10.1007/s10021-018-0255-z

Lair GJ, Zehetner F, Khan ZH, Gerzabek MH (2009) Phosphorus sorption–desorption in alluvial soils of a young weathering sequence at the Danube River. Geoderma 149:39–44. https://doi.org/10.1016/j.geoderma.2008.11.011

Larned ST, Nikora VI, Biggs BJF (2004) Mass-transfer-limited nitrogen and phosphorus uptake by stream periphyton: A conceptual model and experimental evidence. Limnol Oceanogr 49:1992–2000. https://doi.org/10.4319/lo.2004.49.6.1992

Lee L (2020) NADA: Nondetects and Data Analysis for Enivironmental Data. Version 1.6-1.1URL https://CRAN.R-project.org/package=NADA

Lewandowski J, Nützmann G (2010) Nutrient retention and release in a floodplain's aquifer and in the hyporheic zone of a lowland river. Ecol Eng 36:1156–1166. https://doi.org/10.1016/j.ecoleng.2010.01.005

Li A, Bernal S, Kohler B, et al (2021) Residence Time in Hyporheic Bioactive Layers Explains Nitrate Uptake in Streams. Water Resources Research n/a:e2020WR027646. https://doi.org/10.1029/2020WR027646

Li H, Liu Y, Cao X, et al (2016) Functions of Calcium-bound Phosphorus in Relation to Characteristics of Phosphorus Releasing Bacteria in Sediment of a Chinese Shallow Lake (Lake Wabu). Geomicrobiology Journal 33:751–757. https://doi.org/10.1080/01490451.2015.1099762

Liberati A, Altman DG, Tetzlaff J, et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 6:. https://doi.org/10.1371/journal.pmed.1000100

Licht C (2010) New methods for generating significance levels from multiply-imputed data. University of Bamberg

Lijklema L (1980) Interaction of orthophosphate with iron(III) and aluminum hydroxides. Environ Sci Technol 14:537–541. https://doi.org/10.1021/es60165a013

Limousin G, Gaudet JP, Charlet L, et al (2007) Sorption isotherms: A review on physical bases, modeling and measurement. Appl Geochem 22:249–275. https://doi.org/10.1016/j.apgeochem.2006.09.010

Lin C, Wang Z, He M, et al (2009) Phosphorus sorption and fraction characteristics in the upper, middle and low reach sediments of the Daliao river systems, China. Journal of Hazardous Materials 170:278–285. https://doi.org/10.1016/j.jhazmat.2009.04.102

Lindsay WL (1979) Chemical equilibria in soils. John Wiley and Sons Ltd.

Lindsay WL, Vlek PLG, Chien SH (1989) Phosphate Minerals. In: Minerals in Soil Environments, 2nd edn. Soil Science Society of America, pp 1089–1130

Little RJA (1992) Regression with Missing X's: A Review. Journal of the American Statistical Association 87:1227–1237. https://doi.org/10.1080/01621459.1992.10476282

Little RJA, Rubin DB (2002) Statistical Analysis with Missing Data, 2nd edn. John Wiley & Sons, Inc., New York

Loh PS, Molot LA, Nowak E, et al (2013) Evaluating relationships between sediment chemistry and anoxic phosphorus and iron release across three different water bodies. Inland Waters 3:105–118. https://doi.org/10.5268/IW-3.1.533

Long MH, Rheuban JE, Berg P, Zieman JC (2012) A comparison and correction of light intensity loggers to photosynthetically active radiation sensors. Limnology and Oceanography: Methods 10:416–424. https://doi.org/10.4319/lom.2012.10.416

Lottig NR, Stanley EH (2007) Benthic sediment influence on dissolved phosphorus concentrations in a headwater stream. Biogeochemistry 84:297–309. https://doi.org/10.1007/s10533-007-9116-0

Lovley DR (1991) Dissimilatory Fe(III) and Mn(IV) reduction. Microbiology and Molecular Biology Reviews 55:259–287

Lucci GM, McDowell RW, Condron LM (2010) Evaluation of base solutions to determine equilibrium phosphorus concentrations (EPC<inf>0</inf>) in stream sediments. International Agrophysics 24:157–163

Lukkari K, Hartikainen H, Leivuori M (2007) Fractionation of sediment phosphorus revisited. I: Fractionation steps and their biogeochemical basis. Limnol Oceanogr Methods 5:433–444. https://doi.org/10.4319/lom.2007.5.433

Machesky ML, Holm TR, Slowikowski JA (2010) Phosphorus Speciation in Stream Bed Sediments from an Agricultural Watershed: Solid-Phase Associations and Sorption Behavior. Aquat Geochem 16:639–662. https://doi.org/10.1007/s10498-010-9103-2

Macintosh KA, Mayer BK, McDowell RW, et al (2018) Managing Diffuse Phosphorus at the Source versus at the Sink. Environ Sci Technol 52:11995–12009. https://doi.org/10.1021/acs.est.8b01143

Malá J, Lagová M (2014) Comparison of digestion methods for determination of total phosphorus in river sediments. Chem Pap 68:1015–1021. https://doi.org/10.2478/s11696-014-0555-5

Mallakpour I, Villarini G (2015) The changing nature of flooding across the central United States. Nature Clim Change 5:250–254. https://doi.org/10.1038/nclimate2516

Manzoni S, Porporato A (2011) Common hydrologic and biogeochemical controls along the soil–stream continuum. Hydrological Processes 25:1355–1360. https://doi.org/10.1002/hyp.7938

Maranger R, Jones SE, Cotner JB (2018) Stoichiometry of carbon, nitrogen, and phosphorus through the freshwater pipe. Limnology and Oceanography Letters 3:89–101. https://doi.org/10.1002/lol2.10080

Marcarelli AM, Baker MA, Wurtsbaugh WA (2008) Is in-stream N2 fixation an important N source for benthic communities and stream ecosystems? Journal of the North American Benthological Society 27:186–211. https://doi.org/10.1899/07-027.1

Marcé R, Schiller DV, Aguilera R, et al (2018) Contribution of hydrologic opportunity and biogeochemical reactivity to the variability of nutrient retention in river networks. 1–13. https://doi.org/10.1002/2017GB005677

Marti E, Aumatell J, Godé L, et al (2004) Nutrient Retention Efficiency in Streams Receiving Inputs from Wastewater Treatment Plants. Journal of Environmental Quality 33:285–293. https://doi.org/10.2134/jeq2004.2850

Martí E, Feijoó C, Vilches C, et al (2020) Diel variation of nutrient retention is associated with metabolism for ammonium but not phosphorus in a lowland stream. Freshwater Science 39:268–280. https://doi.org/10.1086/708933

Marton JM, Roberts BJ (2014) Spatial variability of phosphorus sorption dynamics in Louisiana salt marshes. J Geophys Res Biogeosci 119:451–465. https://doi.org/10.1002/2013JG002486

Maruo M, Ishimaru M, Azumi Y, et al (2016) Comparison of soluble reactive phosphorus and orthophosphate concentrations in river waters. Limnology 17:7–12. https://doi.org/10.1007/s10201-015-0463-6

Matheson FE, Quinn JM, Martin ML (2012) Effects of irradiance on diel and seasonal patterns of nutrient uptake by stream periphyton. Freshwater Biology 57:1617–1630. https://doi.org/10.1111/j.1365-2427.2012.02822.x

McDaniel MD, David MB, Royer TV (2009) Relationships between benthic sediments and water column phosphorus in Illinois streams. J Environ Qual 38:607–617. https://doi.org/10.2134/jeq2008.0094

McDonnell JJ (2013) Are all runoff processes the same? Hydrological Processes 27:4103–4111. https://doi.org/10.1002/hyp.10076

McDowell RW (2015) Relationship between sediment chemistry, equilibrium phosphorus concentrations, and phosphorus concentrations at baseflow in rivers of the New Zealand National River Water Quality Network. J Environ Qual 44:921–929. https://doi.org/10.2134/jeq2014.08.0362

McDowell RW (2003) Sediment phosphorus chemistry and microbial biomass along a lowland New Zealand stream. Aquat Geochem 9:19–40. https://doi.org/10.1023/B:AQUA.0000005620.15485.6d

McDowell RW, Biggs BJF, Sharpley AN, Nguyen L (2004) Connecting phosphorus loss from agricultural landscapes to surface water quality. Chem Ecol 20:1–40. https://doi.org/10.1080/02757540310001626092

McDowell RW, Cox N, Daughney CJ, et al (2015) A national assessment of the potential linkage between soil, and surface and groundwater concentrations of phosphorus. J Am Water Resour Assoc 51:992–1002. https://doi.org/10.1111/1752-1688.12337

McDowell RW, Depree C, Stenger R (2020a) Likely controls on dissolved reactive phosphorus concentrations in baseflow of an agricultural stream. J Soils Sediments 20:3254–3265. https://doi.org/10.1007/s11368-020-02644-w

McDowell RW, Elkin KR, Kleinman PJA (2017) Temperature and nitrogen effects on phosphorus uptake by agricultural stream-bed sediments. J Environ Qual 46:295–301. https://doi.org/10.2134/jeq2016.09.0352

McDowell RW, Hedley MJ, Pletnyakov P, et al (2019a) Why are median phosphorus concentrations improving in New Zealand streams and rivers? Journal of the Royal Society of New Zealand 0:1–28. https://doi.org/10.1080/03036758.2019.1576213

McDowell RW, Hill SJ (2015) Speciation and distribution of organic phosphorus in river sediments: a national survey. J Soils Sediments 15:2369–2379. https://doi.org/10.1007/s11368-015-1125-3

McDowell RW, Noble A, Pletnyakov P, et al (2020b) Global mapping of freshwater nutrient enrichment and periphyton growth potential. Sci Rep 10:1–13. https://doi.org/10.1038/s41598-020-60279-w

McDowell RW, Schallenberg M, Larned S (2018) A strategy for optimizing catchment management actions to stressor–response relationships in freshwaters. Ecosphere 9:e02482. https://doi.org/10.1002/ecs2.2482

McDowell RW, Sharpley AN (2003a) Phosphorus solubility and release kinetics as a function of soil test P concentration. Geoderma 112:143–154. https://doi.org/10.1016/S0016-7061(02)00301-4

McDowell RW, Sharpley AN (2003b) Uptake and Release of Phosphorus from Overland Flow in a Stream Environment. J Environ Qual 32:937–948. https://doi.org/10.2134/jeq2003.9370

McDowell RW, Sharpley AN, Folmar G (2003) Modification of phosphorus export from an eastern USA catchment by fluvial sediment and phosphorus inputs. Agriculture, Ecosystems and Environment 99:187–199. https://doi.org/10.1016/S0167-8809(03)00142-7

McDowell RW, Simpson ZP, Stenger R, Depree C (2019b) The influence of a flood event on the potential sediment control of baseflow phosphorus concentrations in an intensive agricultural catchment. J Soils Sediments 19:429–438. https://doi.org/10.1007/s11368-018-2063-7

McDowell RW, Snelder TH, Cox N, et al (2013) Establishment of reference or baseline conditions of chemical indicators in New Zealand streams and rivers relative to present conditions. Mar Freshwater Res 64:387–400. https://doi.org/10.1071/MF12153

McLaren AD (1969) Radiation as a technique in soil biology and biochemistry. Soil Biology and Biochemistry 1:63–73. https://doi.org/10.1016/0038-0717(69)90035-2

McLaughlin MJ, Alston AM, Martin JK (1986) Measurement of phosphorus in the soil microbial biomass: A modified procedure for field soils. Soil Biology and Biochemistry 18:437–443. https://doi.org/10.1016/0038-0717(86)90050-7

McNally SR, Beare MH, Curtin D, et al (2017) Soil carbon sequestration potential of permanent pasture and continuous cropping soils in New Zealand. Global Change Biology 23:4544–4555. https://doi.org/10.1111/gcb.13720

McNamara NP, Black HIJ, Beresford NA, Parekh NR (2003) Effects of acute gamma irradiation on chemical, physical and biological properties of soils. Applied Soil Ecology 24:117–132. https://doi.org/10.1016/S0929-1393(03)00073-8

Meals DW, Dressing SA, Davenport TE (2010) Lag time in water quality response to best management practices: a review. J Environ Qual 39:85–96. https://doi.org/10.2134/jeq2009.0108

Mendes LRD, Tonderski K, Kjaergaard C (2018) Phosphorus accumulation and stability in sediments of surface-flow constructed wetlands. Geoderma 331:109–120. https://doi.org/10.1016/j.geoderma.2018.06.015

Meng X-L (1994) Multiple-Imputation Inferences with Uncongenial Sources of Input. Statistical Science 9:538–558

Mengersen K, Gurevitch J, Schmid CH (2013a) Meta-analysis of primary data. In: Handbook of meta-analysis in ecology and evolution. Princeton University Press, pp 300–312

Mengersen K, Schmidt C, Jennions M, et al (2013b) Statistical models and approaches to inference. In: Handbook of Meta-analysis in Ecology and Evolution. pp 89–107

Meyer JL (1994) The microbial loop in flowing waters. Microb Ecol 28:195–199. https://doi.org/10.1007/BF00166808

Meyer JL (1979) The role of sediments and bryophytes in phosphorus dynamics in a head water stream ecosystem. Limnology & Oceanography 24:365–375. https://doi.org/10.4319/lo.1979.24.2.0364

Miller RO (1997) Microwave digestion of plant tissue in a closed vessel. In: Handbook of Reference Methods for Plant Analysis. CRC Press, Boca Raton, pp 69–73

Miltenburg JC, Golterman HL (1998) The energy of the adsorption of o-phosphate onto ferric hydroxide. Hydrobiol 364:93–97. https://doi.org/10.1023/A:1003107907214

Mitchell M, Muftakhidinov B, Winchen T, Trande A (2020) Engauge Digitizer, Version 12. Version 12URL http://markummitchell.github.io/engauge-digitizer/

Monbet P, D. McKelvie I, J. Worsfold P (2010) Sedimentary pools of phosphorus in the eutrophic Tamar estuary (SW England). J Environ Monit 12:296–304. https://doi.org/10.1039/B911429G

Mongeon P, Paul-Hus A (2016) The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics 106:213–228. https://doi.org/10.1007/s11192-015-1765-5

Moriasi DN, Arnold JG, Van Liew MW, et al (2007) Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50:885–900. https://doi.org/10.13031/2013.23153

Mulholland PJ (1996) Role in nutrient cycling in streams. In: Stevenson R, Bothwell ML, Lowe R, Thorp JH (eds) Algal Ecology, 1st edn. Academic Press

Mulholland PJ, Marzolf ER, Webster JR, et al (1997) Evidence that hyporheic zones increase heterotrophic metabolism and phosphorus uptake in forest streams. Limnol Oceanogr 42:443–451. https://doi.org/10.4319/lo.1997.42.3.0443

Mulholland PJ, Newbold JD, Elwood JW, Hom CL (1983) The effect of grazing intensity on phosphorus spiralling in autotropic streams. Oecologia 58:358–366. https://doi.org/10.1007/BF00385236

Mulholland PJ, Steinman AD, Marzolf ER, et al (1994) Effect of periphyton biomass on hydraulic characteristics and nutrient cycling in streams. Oecologia 98:40–47. https://doi.org/10.1007/BF00326088

Munn NL, Meyer JL (1990) Habitat-Specific Solute Retention in Two Small Streams: An Intersite Comparison. Ecology 71:2069–2082. https://doi.org/10.2307/1938621

Murphy J, Riley JP (1962) A modified single solution method for the determination of phosphate in natural waters. Anal Chim Acta 27:31–36. https://doi.org/10.1016/S0003-2670(00)88444-5

Muscarella ME, Bird KC, Larsen ML, et al (2014) Phosphorus resource heterogeneity in microbial food webs. Aquat Microb Ecol 73:259–272

Nagul EA, McKelvie ID, Worsfold P, Kolev SD (2015) The molybdenum blue reaction for the determination of orthophosphate revisited: Opening the black box. Anal Chim Acta 890:60–82. https://doi.org/10.1016/j.aca.2015.07.030

Nair PS, Logan TJ, Sharpley AN, et al (1984) Interlaboratory comparison of a standardized phosphorus adsorption procedure. Journal of Environment Quality 13:591–595

Nakagawa S, Freckleton RP (2011) Model averaging, missing data and multiple imputation: a case study for behavioural ecology. Behav Ecol Sociobiol 65:103–116. https://doi.org/10.1007/s00265-010-1044-7

Nakagawa S, Noble DWA, Senior AM, Lagisz M (2017) Meta-evaluation of meta-analysis: ten appraisal questions for biologists. BMC Biology 15:. https://doi.org/10.1186/s12915-017-0357-7

Neal C, Jarvie HP, Williams RJ, et al (2002) Phosphorus-calcium carbonate saturation relationships in a lowland chalk river impacted by sewage inputs and phosphorus remediation: an assessment of phosphorus self-cleansing mechanisms in natural waters. Sci Total Environ 282–283:295–310. https://doi.org/10.1016/S0048-9697(01)00920-2

Newbold JD, Elwood JW, O'Neill RV, Sheldon AL (1983) Phosphorus dynamics in a woodland stream ecosystem: a study of nutrient spiralling. Ecology 64:1249–1265

Nimick DA, Gammons CH, Parker SR (2011) Diel biogeochemical processes and their effect on the aqueous chemistry of streams: A review. Chem Geol 283:3–17. https://doi.org/10.1016/j.chemgeo.2010.08.017

Ocampo CJ, Oldham CE, Sivapalan M (2006) Nitrate attenuation in agricultural catchments: Shifting balances between transport and reaction. Water Resources Research 42:. https://doi.org/10.1029/2004WR003773

Öhlinger H, Von Mersi W (1996) Enzymes Involved in Intracellular Metabolism. In: Schinner F, Öhlinger R, Kandeler E, Margesin R (eds) Methods in Soil Biology. Springer, Berlin, Heidelberg, pp 235–245

Ohno T, Zibilske LM (1991) Determination of low concentrations of phosphorus in soil extracts using malachite green. Soil Sci Soc Am J 55:892–895

Olsen SR, Sommers LE (1982) Phosphorus. In: Klute A, Page AL (eds) Methods of Soil Analysis Part 2, 2nd edn. American Society of Agronomy, pp 403–430

Orr CH, Clark JJ, Wilcock PR, et al (2009) Comparison of morphological and biological control of exchange with transient storage zones in a field-scale flume. Journal of Geophysical Research: Biogeosciences 114:. https://doi.org/10.1029/2008JG000825

Östlund P, Carman R, Edvardsson UG, Hallstadius L (1989) Sterilization of sediments by ionizing radiation. Applied Geochemistry 4:99–103. https://doi.org/10.1016/0883-2927(89)90062-0

Oviedo-Vargas D, Royer TV, Johnson LT (2013) Dissolved organic carbon manipulation reveals coupled cycling of carbon, nitrogen, and phosphorus in a nitrogen-rich stream. Limnology and Oceanography 58:1196–1206. https://doi.org/10.4319/lo.2013.58.4.1196

Owens PN, Batalla RJ, Collins AJ, et al (2005) Fine-grained sediment in river systems: environmental significance and management issues. River Research and Applications 21:693–717. https://doi.org/10.1002/rra.878

Packman AI, Salehin M (2003) Relative roles of stream flow and sedimentary conditions in controlling hyporheic exchange. Hydrobiol 494:291–297. https://doi.org/10.1023/A:1025403424063

Palmer-Felgate EJ, Bowes MJ, Stratford C, et al (2011) Phosphorus release from sediments in a treatment wetland: Contrast between DET and EPC0 methodologies. Ecol Eng 37:826–832. https://doi.org/10.1016/j.ecoleng.2010.12.024

Palmer-Felgate EJ, Mortimer RJG, Krom MD, Jarvie HP (2010) Impact of point-source pollution on phosphorus and nitrogen cycling in stream-bed sediments. Environ Sci Technol 44:908–914. https://doi.org/10.1021/es902706r

Parfitt RL (1979) Anion Adsorption by Soils and Soil Materials. In: Brady NC (ed) Advances in Agronomy. Academic Press, USA, pp 1–50

Parikh SJ, Goyne KW, Margenot AJ, et al (2014) Chapter One - Soil Chemical Insights Provided through Vibrational Spectroscopy. In: Sparks DL (ed) Advances in Agronomy. Academic Press, pp 1–148

Parker SP, Bowden WB, Flinn MB (2016) The effect of acid strength and postacidification reaction time on the determination of chlorophyll a in ethanol extracts of aquatic periphyton. Limnol Oceanogr Methods 14:839–852. https://doi.org/10.1002/lom3.10130

Parker SP, Bowden WB, Flinn MB, et al (2018) Effect of particle size and heterogeneity on sediment biofilm metabolism and nutrient uptake scaled using two approaches. Ecosphere 9:e02137. https://doi.org/10.1002/ecs2.2137

Parker SR, Gammons CH, Poulson SR, DeGrandpre MD (2007) Diel variations in stream chemistry and isotopic composition of dissolved inorganic carbon, upper Clark Fork River, Montana, USA. Applied Geochemistry 22:1329–1343. https://doi.org/10.1016/j.apgeochem.2007.02.007

Parkhurst DL, Appelo CAJ (2013) Description of input and examples for PHREEQC version 3: a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. U.S. Geological Survey, Reston, VA

Parsons CT, Rezanezhad F, O'Connell DW, Cappellen PV (2017) Sediment phosphorus speciation and mobility under dynamic redox conditions. Biogeosciences 14:3585–3602. https://doi.org/10.5194/bg-14-3585-2017

Peiffer S, Kappler A, Haderlein SB, et al (2021) A biogeochemical–hydrological framework for the role of redox-active compounds in aquatic systems. Nature Geoscience 14:264–272. https://doi.org/10.1038/s41561-021-00742-z

Peryer-Fursdon J, Abell JM, Clarke D, et al (2014) Spatial variability in sediment phosphorus characteristics along a hydrological gradient upstream of Lake Rotorua, New Zealand. Environmental Earth Sciences 73:1573–1585. https://doi.org/10.1007/s12665-014-3508-y

Pettersson K, Boström B, Jacobsen O-S (1988) Phosphorus in sediments — speciation and analysis. Hydrobiol 170:91–101. https://doi.org/10.1007/BF00024900

Phillips EJP, Lovley DR (1987) Determination of Fe(III) and Fe(II) in Oxalate Extracts of Sediment. Soil Science Society of America Journal 51:938–941. https://doi.org/10.2136/sssaj1987.03615995005100040021x

Pierzynski GM (2000) Methods of phosphorus analysis for soils, sediments, residuals, and waters. North Carolina State University

Pinheiro JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. Springer, New York

Pinheiro JC, Bates DM, Debroy S, Sarkar D (2020) nlme: Linear and Nonlinear Mixed Effects Models. Version R package version 3.1-144URL https://CRAN.R-project.org/package=nlme

Piper J (2005) Water Resources of the Reporoa Basin. Environment Waikato Regional Council, Hamilton, New Zealand

Piper LR, Cross WF, McGlynn BL (2017) Colimitation and the coupling of N and P uptake kinetics in oligotrophic mountain streams. Biogeochemistry 132:165–184. https://doi.org/10.1007/s10533-017-0294-0

Plant LJ, House WA (2002) Precipitation of calcite in the presence of inorganic phosphate. Colloids Surf, A 203:143–153. https://doi.org/10.1016/S0927-7757(01)01089-5

Powers SM, Bruulsema TW, Burt TP, et al (2016) Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat Geosci 9:353–356. https://doi.org/10.1038/ngeo2693

Powlson DS, Jenkinson DS (1976) The effects of biocidal treatments on metabolism in soil—II. Gamma irradiation, autoclaving, air-drying and fumigation. Soil Biology and Biochemistry 8:179–188. https://doi.org/10.1016/0038-0717(76)90002-X

Pribyl DW (2010) A critical review of the conventional SOC to SOM conversion factor. Geoderma 156:75–83. https://doi.org/10.1016/j.geoderma.2010.02.003

Prosser JA, Speir TW, Stott DE (2011) Soil Oxidoreductases and FDA Hydrolysis. In: Methods of Soil Enzymology. Soil Science Society of America, pp 103–124

Qiu S, McComb AJ (1995) Planktonic and microbial contributions to phosphorus release from fresh and air-dried sediments. Mar Freshwater Res 46:1039–1045. https://doi.org/10.1071/mf9951039

Qiu S, McComb AJ (2002) Interrelations between iron extractability and phosphate sorption in reflooded air-dried sediments. Hydrobiologia 472:39–44. https://doi.org/10.1023/A:1016317100164

Quinn JM, Cooper AB, Stroud MJ, Burrell GP (1997) Shade effects on stream periphyton and invertebrates: An experiment in streamside channels. New Zealand Journal of Marine and Freshwater Research 31:665–683. https://doi.org/10.1080/00288330.1997.9516797

Quinn JM, Rutherford JC, Schiff SJ (2020) Nutrient attenuation in a shallow, gravel-bed river. I. In-situ chamber experiments. New Zealand Journal of Marine and Freshwater Research 54:393–409. https://doi.org/10.1080/00288330.2020.1740284

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Rahutomo S, Kovar JL, Thompson ML (2018) Varying redox potential affects P release from stream bank sediments. PLOS ONE 13:e0209208. https://doi.org/10.1371/journal.pone.0209208

Rantz SE (1982) Measurement and computation of streamflow: Volume 1, Measurement of stage and discharge. United States Geological Survey, Washington, D.C.

Rawlins BG (2011) Controls on the phosphorus content of fine stream bed sediments in agricultural headwater catchments at the landscape-scale. Agriculture, Ecosystems & Environment 144:352–363. https://doi.org/10.1016/j.agee.2011.10.002

Raymond PA, Zappa CJ, Butman D, et al (2012) Scaling the gas transfer velocity and hydraulic geometry in streams and small rivers. Limnology and Oceanography 41–53. https://doi.org/10.1215/21573689-1597669@10.1002/(ISSN)1939-5590.MethaneVI

Records RM, Wohl E, Arabi M (2016) Phosphorus in the river corridor. Earth-Science Reviews 158:65–88. https://doi.org/10.1016/j.earscirev.2016.04.010

Reddy KR, Kadlec RH, Flaig E, Gale PM (1999) Phosphorus Retention in Streams and Wetlands: A Review. Critical Reviews in Environmental Science and Technology 29:83–146. https://doi.org/10.1080/10643389991259182

Reeder WJ, Quick AM, Farrell TB, et al (2018) Spatial and Temporal Dynamics of Dissolved Oxygen Concentrations and Bioactivity in the Hyporheic Zone. Water Resources Research 54:2112–2128. https://doi.org/10.1002/2017WR021388

Ren J, Packman AI (2005) Coupled stream-subsurface exchange of colloidal hematite and dissolved zinc, copper, and phosphate. Environ Sci Technol 39:6387–6394. https://doi.org/10.1021/es050168q

Rier ST, Kuehn KA, Francoeur SN (2007) Algal regulation of extracellular enzyme activity in stream microbial communities associated with inert substrata and detritus. Journal of the North American Benthological Society 26:439–449. https://doi.org/10.1899/06-080.1

Rier ST, Shirvinski JM, Kinek KC (2014) In situ light and phosphorus manipulations reveal potential role of biofilm algae in enhancing enzyme-mediated decomposition of organic matter in streams. Freshwater Biol 59:1039–1051. https://doi.org/10.1111/fwb.12327

Rier ST, Stevenson RJ (2001) Relation of environmental factors to density of epilithic lotic bacteria in 2 ecoregions. Journal of the North American Benthological Society 20:520–532. https://doi.org/10.2307/1468085

Rietra RPJJ, Hiemstra T, Van Riemsdijk WH (2001) Interaction between calcium and phosphate adsorption on goethite. Environmental Science and Technology 35:3369–3374. https://doi.org/10.1021/es000210b

Rinaldo A, Benettin P, Harman CJ, et al (2015) Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes. Water Resources Research 51:4840–4847. https://doi.org/10.1002/2015WR017273

Roberts BJ, Mulholland PJ, Hill WR (2007) Multiple Scales of Temporal Variability in Ecosystem Metabolism Rates: Results from 2 Years of Continuous Monitoring in a Forested Headwater Stream. Ecosystems 10:588–606. https://doi.org/10.1007/s10021-007-9059-2

Roberts EJ, Cooper RJ (2018) Riverbed sediments buffer phosphorus concentrations downstream of sewage treatment works across the River Wensum catchment, UK. J Soils Sediments 18:2107–2116. https://doi.org/10.1007/s11368-018-1939-x

Rosenberg MS, Rothstein HR, Gurevitch J (2013) Effect sizes: conventional choices and calculations. In: Handbook of Meta-analysis in Ecology and Evolution. pp 61–71

Rothe M, Frederichs T, Eder M, et al (2014) Evidence for vivianite formation and its contribution to long-term phosphorus retention in a recent lake sediment: a novel analytical approach. Biogeosciences 11:5169–5180. https://doi.org/10.5194/bg-11-5169-2014

Rounds SA (2012) Alkalinity and acid neutralizing capacity. In: Wilde FD, Radtke DB (eds) National field manual for the collection of water-quality data. U.S. Geological Survey, Reston, Virginia, pp 1–45

Rubin DB (1996) Multiple Imputation after 18+ Years. Journal of the American Statistical Association 91:473–489. https://doi.org/10.1080/01621459.1996.10476908

Rubin DB (1987) Multiple imputation for nonresponse in surveys. John Wiley & Sons, New York

Rubin DB (1976) Inference and missing data. Biometrika 63:581–592. https://doi.org/10.1093/biomet/63.3.581

Runkel RL, Kimball BA, McKnight DM, Bencala KE (1999) Reactive solute transport in streams: A surface complexation approach for trace metal sorption. Water Resour Res 35:3829–3840. https://doi.org/10.1029/1999WR900259

Rutherford, PM, McGill WB, Arocena JM (2008) Ch 22: Total nitrogen. In: Soil Sampling and Methods of Analysis, 2nd edn. CRC Press, Boca Raton, FL, p 12

Ruttenberg KC (1992) Development of a sequential extraction method for different forms of phosphorus in marine sediments. Limnology and Oceanography 37:1460–1482. https://doi.org/10.4319/lo.1992.37.7.1460

Ryan RJ, Packman AI, Kilham SS (2007) Relating phosphorus uptake to changes in transient storage and streambed sediment characteristics in headwater tributaries of Valley Creek, an urbanizing watershed. J Hydrol 336:444–457. https://doi.org/10.1016/j.jhydrol.2007.01.021

Ryden JC, Syers JK (1975) Charge relationships of phosphate sorption. Nature 255:51–53

Sauer D, Saccone L, Conley DJ, et al (2006) Review of methodologies for extracting plant-available and amorphous Si from soils and aquatic sediments. Biogeochemistry 80:89–108. https://doi.org/10.1007/s10533-005-5879-3

Saunders WMH (1965) Phosphate retention by New Zealand soils and its relationship to free sesquioxides, organic matter, and other soil properties. N Z J Agric Res 8:30–57. https://doi.org/10.1080/00288233.1965.10420021

Schafer JL, Olsen MK (1998) Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective. Multivariate Behavioral Research 33:545–571. https://doi.org/10.1207/s15327906mbr3304_5

Schlesinger WH, Bernhardt ES (2013) Chapter 4 - The Lithosphere. In: Schlesinger WH, Bernhardt ES (eds) Biogeochemistry (Third Edition). Academic Press, Boston, pp 93–133

Schlichting A, Leinweber P (2002) Effects of pretreatment on sequentially-extracted phosphorus fractions from peat soils. Communications in Soil Science and Plant Analysis 33:1617–1627. https://doi.org/10.1081/CSS-120004303

Scott JT, Cotner J, LaPara T (2012) Variable Stoichiometry and Homeostatic Regulation of Bacterial Biomass Elemental Composition. Front Microbiol 3:. https://doi.org/10.3389/fmicb.2012.00042

Scott JT, Haggard BE, Sharpley AN, Romeis JJ (2011) Change Point Analysis of Phosphorus Trends in the Illinois River (Oklahoma) Demonstrates the Effects of Watershed Management. Journal of Environmental Quality 40:1249–1256. https://doi.org/10.2134/jeq2010.0476

Senn A-C, Kaegi R, Hug SJ, et al (2015) Composition and structure of Fe(III)-precipitates formed by Fe(II) oxidation in water at near-neutral pH: Interdependent effects of phosphate, silicate and Ca. Geochim Cosmochim Acta 162:220–246. https://doi.org/10.1016/j.gca.2015.04.032

Shah AD, Bartlett JW, Carpenter J, et al (2014) Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. Am J Epidemiol 179:764–774. https://doi.org/10.1093/aje/kwt312

Sharpley A, Jarvie HP, Buda A, et al (2013) Phosphorus legacy: Overcoming the effects of past management practices to mitigate future water quality impairment. J Environ Qual 42:1308–1326. https://doi.org/10.2134/jeq2013.03.0098

Sharpley AN (1993) An Innovative Approach to Estimate Bioavailable Phosphorus in Agricultural Runoff Using Iron Oxide-Impregnated Paper. J Environ Qual 22:597–601. https://doi.org/10.2134/jeq1993.00472425002200030026x

Shmueli G (2010) To Explain or to Predict? Statist Sci 25:289–310. https://doi.org/10.1214/10-STS330

Sigg L, Stumm W (1981) The interaction of anions and weak acids with the hydrous goethite (α-FeOOH) surface. Colloids Surf 2:101–117. https://doi.org/10.1016/0166-6622(81)80001-7

Simmons JA (2010) Phosphorus removal by sediment in streams contaminated with acid mine drainage. Water, Air, and Soil Pollution 209:123–132. https://doi.org/10.1007/s11270-009-0185-7

Simpson ZP, McDowell RW, Condron LM (2020) The biotic contribution to the benthic stream sediment phosphorus buffer. Biogeochemistry 151:63–79. https://doi.org/10.1007/s10533-020-00709-z

Simpson ZP, McDowell RW, Condron LM (2019) The error in stream sediment phosphorus fractionation and sorption properties effected by drying pretreatments. J Soils Sediments 19:1587–1597. https://doi.org/10.1007/s11368-018-2180-3

Simpson ZP, McDowell RW, Condron LM, et al (2021) Sediment phosphorus buffering in streams at baseflow: A meta-analysis. Journal of Environmental Quality n/a: https://doi.org/10.1002/jeq2.20202

Sinsabaugh RL, Follstad Shah JJ, Hill BH, Elonen CM (2012) Ecoenzymatic stoichiometry of stream sediments with comparison to terrestrial soils. Biogeochemistry 111:455–467. https://doi.org/10.1007/s10533-011-9676-x

Sinsabaugh RL, Hill BH, Shah JJF (2009) Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment. Nature 462:795–798. https://doi.org/10.1038/nature08632

Small GE, Ardón M, Duff JH, et al (2016) Phosphorus retention in a lowland Neotropical stream following an eight-year enrichment experiment. Freshwater Science 35:1–11. https://doi.org/10.1086/684491

Smil V (2000) Phosphorus in the Environment: Natural Flows and Human Interferences. Annual Review of Energy and the Environment 25:53–88. https://doi.org/10.1146/annurev.energy.25.1.53

Smith DR (2009) Assessment of in-stream phosphorus dynamics in agricultural drainage ditches. Sci Total Environ 407:3883–3889. https://doi.org/10.1016/j.scitotenv.2009.02.038

Smith DR, Warnemuende EA, Haggard BE, Huang C (2006) Dredging of drainage ditches increases short-term transport of soluble phosphorus. Journal of Environmental Quality 35:611–616. https://doi.org/10.2134/jeq2005.0301

Smolders E, Baetens E, Verbeeck M, et al (2017) Internal loading and redox cycling of sediment iron explain reactive phosphorus concentrations in lowland rivers. Environ Sci Technol 51:2584–2592. https://doi.org/10.1021/acs.est.6b04337

Snelder TH, Biggs BJF (2002) Multiscale river environment classification for water resources management. J Am Water Resour Assoc 38:1225–1239. https://doi.org/10.1111/j.1752-1688.2002.tb04344.x

Snelder TH, Biggs BJF, Weatherhead M (2010) New Zealand River Environment Classification User Guide. Ministry for the Environment, Wellington, New Zealand

Sø HU, Postma D, Jakobsen R, Larsen F (2011) Sorption of phosphate onto calcite; results from batch experiments and surface complexation modeling. Geochim Cosmochim Acta 75:2911–2923. https://doi.org/10.1016/j.gca.2011.02.031

Son J-H, Kim S, Carlson KH (2015) Effects of Wildfire on River Water Quality and Riverbed Sediment Phosphorus. Water Air Soil Pollut 226:26. https://doi.org/10.1007/s11270-014-2269-2

Sperazza M, Moore JN, Hendrix MS (2004) High-Resolution Particle Size Analysis of Naturally Occurring Very Fine-Grained Sediment Through Laser Diffractometry. Journal of Sedimentary Research 74:736–743. https://doi.org/10.1306/031104740736

Sposito G (2004) The Surface Chemistry of Natural Particles. Oxford University Press, Oxford, New York

Sprenger M, Stumpp C, Weiler M, et al (2019) The Demographics of Water: A Review of Water Ages in the Critical Zone. Reviews of Geophysics 57:800–834. https://doi.org/10.1029/2018RG000633

Stelzer RS, Heffernan J, Likens GE (2003) The influence of dissolved nutrients and particulate organic matter quality on microbial respiration and biomass in a forest stream. Freshwater Biology 48:1925–1937. https://doi.org/10.1046/j.1365-2427.2003.01141.x

Sterner RW, Elser JJ (2002) Ecological stoichiometry: the biology of elements from molecules to the biosphere. Princeton University Press

Stets EG, Butman D, McDonald CP, et al (2017) Carbonate buffering and metabolic controls on carbon dioxide in rivers. Global Biogeochem Cycles 31:663–677. https://doi.org/10.1002/2016GB005578

Stone M, Mulamoottil G, Logan L (1995) Grain size distribution effects on phosphate sorption by fluvial sediment: implications for modelling sediment-phosphate transport. Hydrological Sciences Journal 40:67–81. https://doi.org/10.1080/02626669509491391

Stookey LL (1970) Ferrozine - a New Spectrophotometric Reagent for Iron. Anal Chem 42:779-. https://doi.org/10.1021/ac60289a016

Strauss R, Brümmer GW, Barrow NJ (1997) Effects of crystallinity of goethite: II. Rates of sorption and desorption of phosphate. Eur J Soil Sci 48:101–114. https://doi.org/10.1111/j.1365-2389.1997.tb00189.x

Stream Solute Workshop (1990) Concepts and Methods for Assessing Solute Dynamics in Stream Ecosystems. Journal of the North American Benthological Society 9:95–119. https://doi.org/10.2307/1467445

Stumm W, Morgan JJ (1996) Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters, Third. John Wiley & Sons, New York

Stumm W, Sulzberger B (1992) The cycling of iron in natural environments: Considerations based on laboratory studies of heterogeneous redox processes. Geochim Cosmochim Acta 56:3233–3257. https://doi.org/10.1016/0016-7037(92)90301-X

Stutter MI, Demars BOLL, Langan SJ (2010) River phosphorus cycling: Separating biotic and abiotic uptake during short-term changes in sewage effluent loading. Water Res 44:4425–4436. https://doi.org/10.1016/j.watres.2010.06.014

Stutter MI, Lumsdon DG (2008) Interactions of land use and dynamic river conditions on sorption equilibria between benthic sediments and river soluble reactive phosphorus concentrations. Water Res 42:4249–4260. https://doi.org/10.1016/j.watres.2008.06.017

Tabatabai MA (1994) Soil Enzymes. In: Weaver RW (ed) Methods of Soil Analysis: Part 2—Microbiological and Biochemical Properties. Soil Science Society of America, pp 775–833

Tang X, Wu M, Dai X, Chai P (2014) Phosphorus storage dynamics and adsorption characteristics for sediment from a drinking water source reservoir and its relation with sediment compositions. Ecol Eng 64:276–284. https://doi.org/10.1016/j.ecoleng.2014.01.005

Tank JL, Rosi-Marshall EJ, Griffiths NA, et al (2010) A review of allochthonous organic matter dynamics and metabolism in streams. J N Am Benthol Soc 29:118–146. https://doi.org/10.1899/08-170.1

Taylor AW, Kunishi HM (1971) Phosphate equilibria on stream sediment and soil in a watershed draining an agricultural region. Journal of Agricultural and Food Chemistry 19:827–831

Trevors JT (1996) Sterilization and inhibition of microbial activity in soil. Journal of Microbiological Methods 26:53–59. https://doi.org/10.1016/0167-7012(96)00843-3

Triska F, Pringle CM, Duff JH, et al (2006) Soluble reactive phosphorus (SRP) transport and retention in tropical, rain forest streams draining a volcanic landscape in Costa Rica: in situ SRP amendment to streams and laboratory studies. Biogeochemistry 81:145–157. https://doi.org/10.1007/s10533-006-9048-0

Turner BL, Haygarth PM (2001) Phosphorus solubilization in rewetted soils. Nature 411:258–258. https://doi.org/10.1038/35077146

Turner BL, Lambers H, Condron LM, et al (2013) Soil microbial biomass and the fate of phosphorus during long-term ecosystem development. Plant Soil 367:225–234. https://doi.org/10.1007/s11104-012-1493-z

Turner BL, Newman S, Cheesman AW, Reddy KR (2007) Sample Pretreatment and Phosphorus Speciation in Wetland Soils. Soil Science Society of America Journal 71:1538–1546. https://doi.org/10.2136/sssaj2007.0017

Turner BL, Romero TE (2010) Stability of hydrolytic enzyme activity and microbial phosphorus during storage of tropical rain forest soils. Soil Biology and Biochemistry 42:459–465. https://doi.org/10.1016/j.soilbio.2009.11.029

Twinch AJ (1987) Phosphate exchange characteristics of wet and dried sediment samples from a hypertrophic reservoir: Implications for the measurements of sediment phosphorus status. Water Research 21:1225–1230. https://doi.org/10.1016/0043-1354(87)90174-6

Tye AM, Rawlins BG, Rushton JC, Price R (2016) Understanding the controls on sediment-P interactions and dynamics along a non-tidal river system in a rural-urban catchment: The River Nene. Appl Geochem 66:219–233. https://doi.org/10.1016/j.apgeochem.2015.12.014

Ugarte MD, Militino AF, Arnholt AT (2015) Probability and Statistics with R, 2nd edn. CRC Press

US EPA O (2019) EPA Method 3050B: Acid Digestion of Sediments, Sludges, and Soils. In: US EPA. https://www.epa.gov/esam/epa-method-3050b-acid-digestion-sediments-sludges-and-soils. Accessed 10 Aug 2020

USEPA (1978) Method 365.3. All Forms of Phosphorus. Methods of Chemical Analysis of Water and Wastes. U.S. Environmental Protection Agency, Cincinatti, OH

Van Buuren S (2018) Flexible Imputation of Missing Data, 2nd edn. Chapman & Hall/CRC, Boca Raton, FL

Van Buuren S, Brand JPL, Groothuis-Oudshoorn CGM, Rubin DB (2006) Fully conditional specification in multivariate imputation. Journal of Statistical Computation and Simulation 76:1049–1064. https://doi.org/10.1080/10629360600810434

Van Buuren S, Groothuis-Oudshoorn K (2011) mice: Multivariate Imputation by Chained Equations in R. J Stat Software 45:1–67. https://doi.org/10.18637/jss.v045.i03

van der Grift B, Rozemeijer JC, Griffioen J, van der Velde Y (2014) Iron oxidation kinetics and phosphate immobilization along the flow-path from groundwater into surface water. Hydrol Earth Syst Sci 18:4687–4702. https://doi.org/10.5194/hess-18-4687-2014

van der Perk M (1997) Effect of Model Structure on the Accuracy and Uncertainty of Results from Water Quality Models. Hydrol Processes 11:227–239. https://doi.org/10.1002/(SICI)1099-1085(19970315)11:3<227::AID-HYP440>3.0.CO;2-#

van der Zee S, Leus F, Louer M (1989) Prediction of phosphate transport in small columns with an approximate sorption kinetics model. Water Resources Research 25:1353–1365. https://doi.org/10.1029/WR025i006p01353

Vaughan MCH, Bowden WB, Shanley JB, et al (2017) High-frequency dissolved organic carbon and nitrate measurements reveal differences in storm hysteresis and loading in relation to land cover and seasonality. Water Resources Research 53:5345–5363. https://doi.org/10.1002/2017WR020491

Vaughan MCH, Bowden WB, Shanley JB, et al (2018) Using in situ UV-Visible spectrophotometer sensors to quantify riverine phosphorus partitioning and concentration at a high frequency. Limnology and Oceanography: Methods 16:840–855. https://doi.org/10.1002/lom3.10287

Venables WN, Ripley BD (2002) Modern applied statistics with S-PLUS, Fourth. Springer, New York

Viollier E, Inglett PW, Hunter K, et al (2000) The ferrozine method revisited: Fe (II)/Fe (III) determination in natural waters. Appl Geochem 15:785–790. https://doi.org/10.1016/S0883-2927(99)00097-9

Wagner BJ, Harvey JW (1997) Experimental design for estimating parameters of rate-limited mass transfer: Analysis of stream tracer studies. Water Resour Res 33:1731–1741. https://doi.org/10.1029/97WR01067

Wang C, Zhang Y, Li H, Morrison RJ (2013) Sequential extraction procedures for the determination of phosphorus forms in sediment. Limnology 14:147–157. https://doi.org/10.1007/s10201-012-0397-1

Wang S, Jin X, Zhao H, Wu F (2006) Phosphorus fractions and its release in the sediments from the shallow lakes in the middle and lower reaches of Yangtze River area in China. Colloids Surf, A 273:109–116. https://doi.org/10.1016/j.colsurfa.2005.08.015

Ward AS, Packman AI (2019) Advancing our predictive understanding of river corridor exchange. Wiley Interdisciplinary Reviews: Water 6:e1327. https://doi.org/10.1002/wat2.1327

Weigelhofer G (2017) The potential of agricultural headwater streams to retain soluble reactive phosphorus. Hydrobiol 793:149–160. https://doi.org/10.1007/s10750-016-2789-4

Weigelhofer G, Ramião JP, Pitzl B, et al (2018a) Decoupled water-sediment interactions restrict the phosphorus buffer mechanism in agricultural streams. Sci Total Environ 628–629:44–52. https://doi.org/10.1016/j.scitotenv.2018.02.030

Weigelhofer G, Ramião JP, Puritscher A, Hein T (2018b) How do chronic nutrient loading and the duration of nutrient pulses affect nutrient uptake in headwater streams? Biogeochemistry 141:249–263. https://doi.org/10.1007/s10533-018-0518-y

Welti N, Striebel M, Ulseth AJ, et al (2017) Bridging Food Webs, Ecosystem Metabolism, and Biogeochemistry Using Ecological Stoichiometry Theory. Front Microbiol 8:. https://doi.org/10.3389/fmicb.2017.01298

Westra S, Fowler HJ, Evans JP, et al (2014) Future changes to the intensity and frequency of short-duration extreme rainfall. Reviews of Geophysics 52:522–555. https://doi.org/10.1002/2014RG000464

Wetzel RG (2001) Limnology, 3rd edn. Academic Press, San Diego, California

White IR, Royston P, Wood AM (2011) Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine 30:377–399. https://doi.org/10.1002/sim.4067

White MJ, Storm DE, Mittelstet A, et al (2014) Development and Testing of an In-Stream Phosphorus Cycling Model for the Soil and Water Assessment Tool. Journal of Environmental Quality 43:215–223. https://doi.org/10.2134/jeq2011.0348

White RE, Beckett PHT (1964) Studies on the phosphate potentials of soils, Part I - The measurement of phosphate potential. Plant Soil 20:1–16. https://doi.org/10.1007/BF01378093

Wilcock RJ, McDowell RW, Quinn JM, et al (2020) Dynamics of phosphorus exchange between sediment and water in a gravel-bed river. New Zealand Journal of Marine and Freshwater Research. https://doi.org/10.1080/00288330.2020.1741402

Wilcock RJ, Scarsbrook MR, Costley KJ, Nagels JW (2002) Controlled release experiments to determine the effects of shade and plants on nutrient retention in a lowland stream. Hydrobiologia 485:153–162. https://doi.org/10.1023/A:1021375509662

Withers PJA, Jarvie HP (2008) Delivery and cycling of phosphorus in rivers: A review. Sci Total Environ 400:379–395. https://doi.org/10.1016/j.scitotenv.2008.08.002

Wohl E (2015) Legacy effects on sediments in river corridors. Earth Sci Rev 147:30–53. https://doi.org/10.1016/j.earscirev.2015.05.001

Wolf DC, Skipper HD (1994) Soil Sterilization. Methods of Soil Analysis: Part 2—Microbiological and Biochemical Properties sssabookseries:41–51. https://doi.org/10.2136/sssabookser5.2.c3

Wood SA, Depree C, Brown L, et al (2015) Entrapped sediments as a source of phosphorus in epilithic cyanobacterial proliferations in low nutrient rivers. PLOS ONE 10:e0141063. https://doi.org/10.1371/journal.pone.0141063

Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73:3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wood SN (2013a) A simple test for random effects in regression models. Biometrika 100:1005–1010. https://doi.org/10.1093/biomet/ast038

Wood SN (2017) Generalized Additive Models : An Introduction with R, Second Edition, Second. Chapman and Hall/CRC

Wood SN (2013b) On p-values for smooth components of an extended generalized additive model. Biometrika 100:221–228. https://doi.org/10.1093/biomet/ass048

Worsfold P, McKelvie I, Monbet P (2016) Determination of phosphorus in natural waters: A historical review. Anal Chim Acta 918:8–20. https://doi.org/10.1016/j.aca.2016.02.047

Worsfold PJ, Gimbert LJ, Mankasingh U, et al (2005) Sampling, sample treatment and quality assurance issues for the determination of phosphorus species in natural waters and soils. Talanta 66:273–293. https://doi.org/10.1016/j.talanta.2004.09.006

Worsfold PJ, Monbet P, Tappin AD, et al (2008) Characterisation and quantification of organic phosphorus and organic nitrogen components in aquatic systems: A Review. Analytica Chimica Acta 624:37–58. https://doi.org/10.1016/j.aca.2008.06.016

Wurtsbaugh WA, Paerl HW, Dodds WK (2019) Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. WIREs Water 6:e1373. https://doi.org/10.1002/wat2.1373

Wymore AS, Potter J, Rodríguez-Cardona B, McDowell WH (2018) Using In-Situ Optical Sensors to Understand the Biogeochemistry of Dissolved Organic Matter Across a Stream Network. Water Resources Research 54:2949–2958. https://doi.org/10.1002/2017WR022168

Zak D, Kleeberg A, Hupfer M (2006) Sulphate-mediated phosphorus mobilization in riverine sediments at increasing sulphate concentration, River Spree, NE Germany. Biogeochemistry 80:109–119. https://doi.org/10.1007/s10533-006-0003-x

Zarnetske JP, Haggerty R, Wondzell SM, Baker MA (2011) Dynamics of nitrate production and removal as a function of residence time in the hyporheic zone. J Geophys Res Biogeosci 116:. https://doi.org/10.1029/2010JG001356

Zhang JZ, Huang XL (2007) Relative importance of solid-phase phosphorus and iron on the sorption behavior of sediments. Environ Sci Technol 41:2789–2795. https://doi.org/10.1021/es061836q

Zhou A, Tang H, Wang D (2005) Phosphorus adsorption on natural sediments: Modeling and effects of pH and sediment composition. Water Res 39:1245–1254. https://doi.org/10.1016/j.watres.2005.01.026