S1 Details on sediment phosphorus and iron fractionation

S1.1 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 H_2SO_4 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).

S1.2 Phosphorus colorimetry

We determined reactive P (RP) in the H2O fraction via malachite-green (detection limit (DL) of 0.006 mg P L^{-1}; D'Angelo et al., 2001) and in the HCl fraction 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. 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; Zhang et al., 1999). The TP values were still valid since digestion removes silicate interference (Malá and Lagová, 2014; Zhang et al., 1999). We recommend that future studies carefully consider silicate interferences and accordingly choose robust methods (Nagul et al., 2015).

S1.3 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). Multiple preliminary tests with Fe(III) standards ($FeCl_3$ in 0.01 M HCl) and Fe-spiked samples indicated incomplete reduction with 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.

S1.4 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. Indeed, the internal reference alerted us to the issue of a change in dithionite chemical. The dithionite chemical in the first two batches of P fractionation was monohydrous and was likely expired; the dithionite chemical in the following fractionations was new, anhydrous, and showed significantly greater extraction efficiency in the BD fractions for the internal reference (20% and 80% increase in P and Fe extraction for BD-I). Therefore, the initial batches of P fractionation were repeated with the better dithionite 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.

S2 Linear models

As there is little fundamental understanding of what sediment variables link to in-stream dissolved reactive P (DRP) concentrations (and how), we fitted linear models for (1) DRP and (2) for sorption metrics (anion storage capacity (ASC) only, as Bache-Williams index had similar results). Our ultimate goal was to explain how sediments influence DRP, and so we used 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.

References

Anastácio, A. S., Harris, B., Yoo, H. I., Fabris, J. D. and Stucki, J. W.: Limitations of the ferrozine method for quantitative assay of mineral systems for ferrous and total iron, Geochim. Cosmochim. Acta, 72(20), 5001–5008, doi:10.1016/j.gca.2008.07.009, 2008.

D'Angelo, E., Crutchfield, J. and Vandiviere, M.: Rapid, sensitive, microscale determination of phosphate in water and soil, J. Environ. Qual., 30(6), 2206–2209, doi:10.2134/jeq2001.2206, 2001.

He, Z. and Honeycutt, C. W.: A modified molybdenum blue method for orthophosphate determination suitable for investigating enzymatic hydrolysis of organic phosphates, Commun. Soil Sci. Plant Anal., 36(9–10), 1373–1383, doi:10.1081/CSS-200056954, 2005.

Jensen, H. S. and Thamdrup, B.: Iron-bound phosphorus in marine sediments as measured by bicarbonate-dithionite extraction, in Proceedings of the Third International Workshop on Phosphorus in Sediments, edited by P. C. M. Boers, Th. E. Cappenberg, and W. van Raaphorst, pp. 47–59, Springer Netherlands., 1993.

Lindsay, W. L.: Chemical equilibria in soils., John Wiley and Sons Ltd., 1979.

Lukkari, K., Hartikainen, H. and Leivuori, M.: Fractionation of sediment phosphorus revisited. I: Fractionation steps and their biogeochemical basis, Limnol. Oceanogr. Methods, 5(12), 433–444, doi:10.4319/lom.2007.5.433, 2007.

Malá, J. and Lagová, M.: Comparison of digestion methods for determination of total phosphorus in river sediments, Chem. Pap., 68(8), 1015–1021, doi:10.2478/s11696-014-0555-5, 2014.

Murphy, J. and Riley, J. P.: A modified single solution method for the determination of phosphate in natural waters, Anal. Chim. Acta, 27, 31–36, doi:10.1016/S0003-2670(00)88444-5, 1962.

Nagul, E. A., McKelvie, I. D., Worsfold, P. and Kolev, S. D.: The molybdenum blue reaction for the determination of orthophosphate revisited: Opening the black box, Anal. Chim. Acta, 890, 60–82, doi:10.1016/j.aca.2015.07.030, 2015.

Sauer, D., Saccone, L., Conley, D. J., Herrmann, L. and Sommer, M.: Review of methodologies for extracting plant-available and amorphous Si from soils and aquatic sediments, Biogeochemistry, 80(1), 89–108, doi:10.1007/s10533-005-5879-3, 2006.

Shmueli, G.: To Explain or to Predict?, Statist. Sci., 25(3), 289–310, doi:10.1214/10-STS330, 2010.

Snelder, T. H., Biggs, B. J. F. and Weatherhead, M.: New Zealand River Environment Classification User Guide, Ministry for the Environment, Wellington, New Zealand., 2010.

Stookey, L. L.: Ferrozine - a New Spectrophotometric Reagent for Iron, Anal. Chem., 42(7), 779–, doi:10.1021/ac60289a016, 1970.

Venables, W. N. and Ripley, B. D.: Modern applied statistics with S-PLUS, Fourth., Springer, New York. [online] Available from: http://www.stats.ox.ac.uk/pub/MASS4, 2002.

Viollier, E., Inglett, P. W., Hunter, K., Roychoudhury, a N. and Van Cappellen, P.: The ferrozine method revisited: Fe (II)/Fe (III) determination in natural waters, Appl. Geochem., 15(6), 785–790, doi:10.1016/S0883-2927(99)00097-9, 2000.

Zhang, J. Z., Fischer, C. J. and Ortner, P. B.: Optimization of performance and minimization of silicate interference in continuous flow phosphate analysis, Talanta, 49(2), 293–304, doi:10.1016/S0039-9140(98)00377-4, 1999.

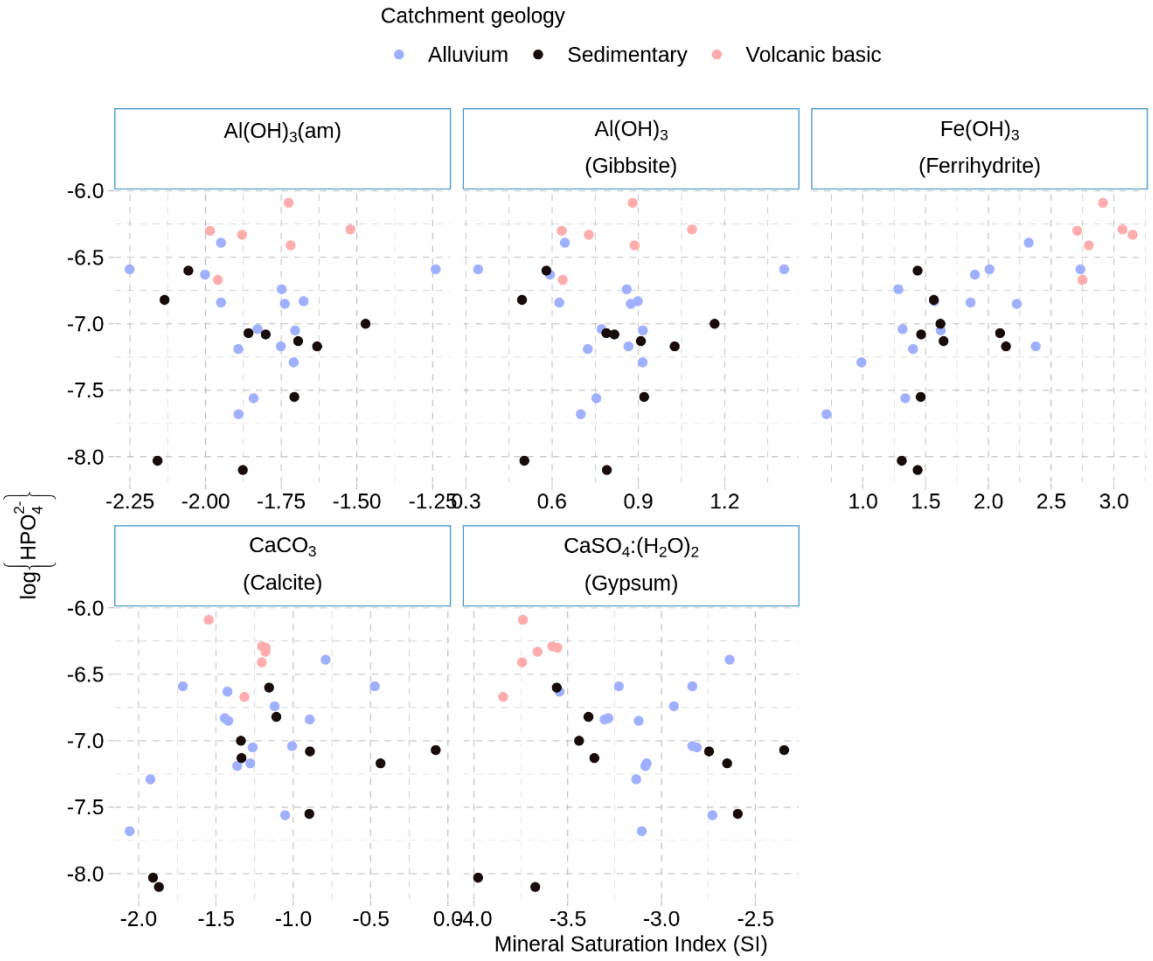


Figure S1. Log-activity of HPO_4^{2-} as a function of select mineral saturation indices (SIs). Mineral formulas are shown as given in the MINTEQA2 v4 database.

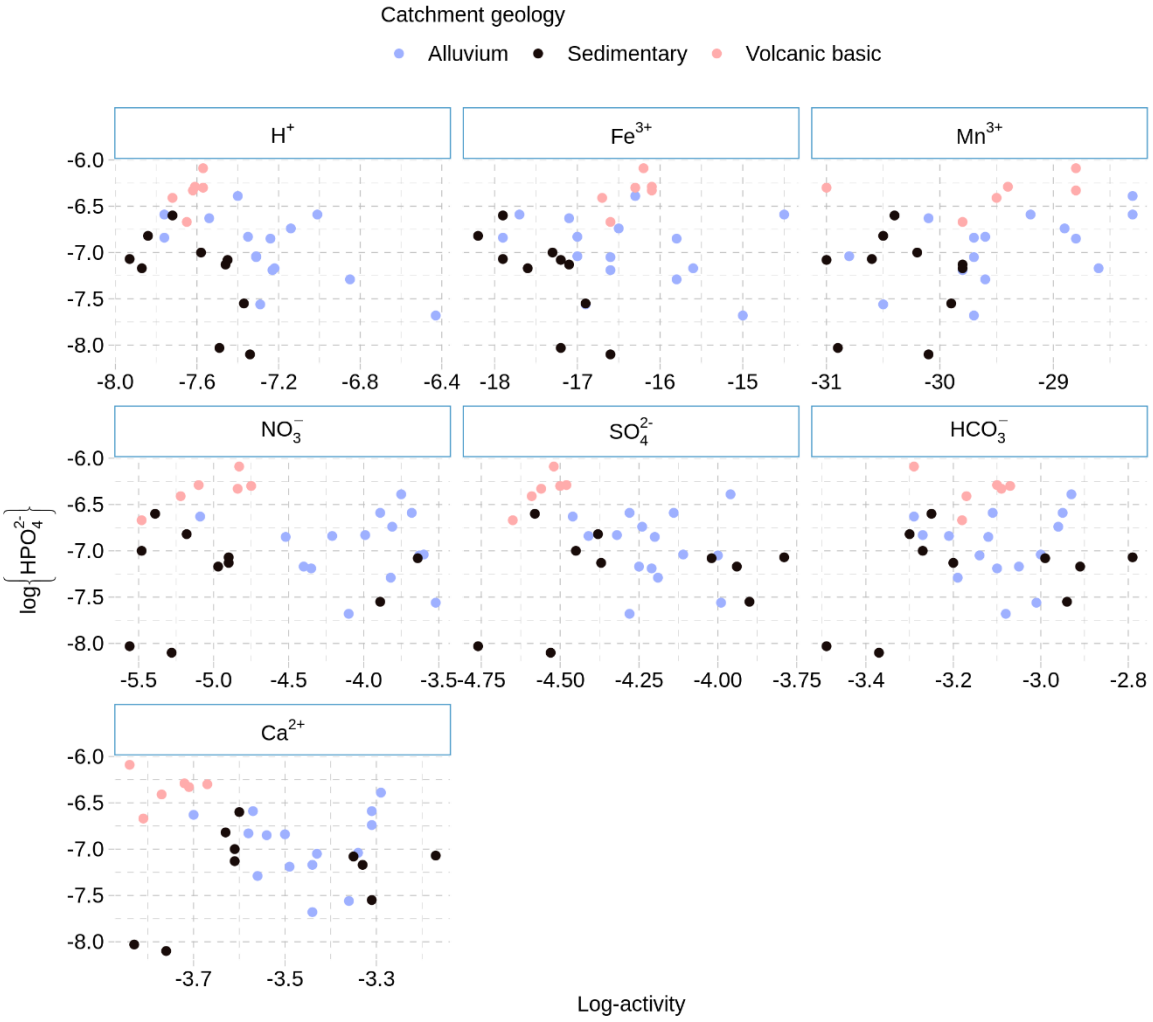
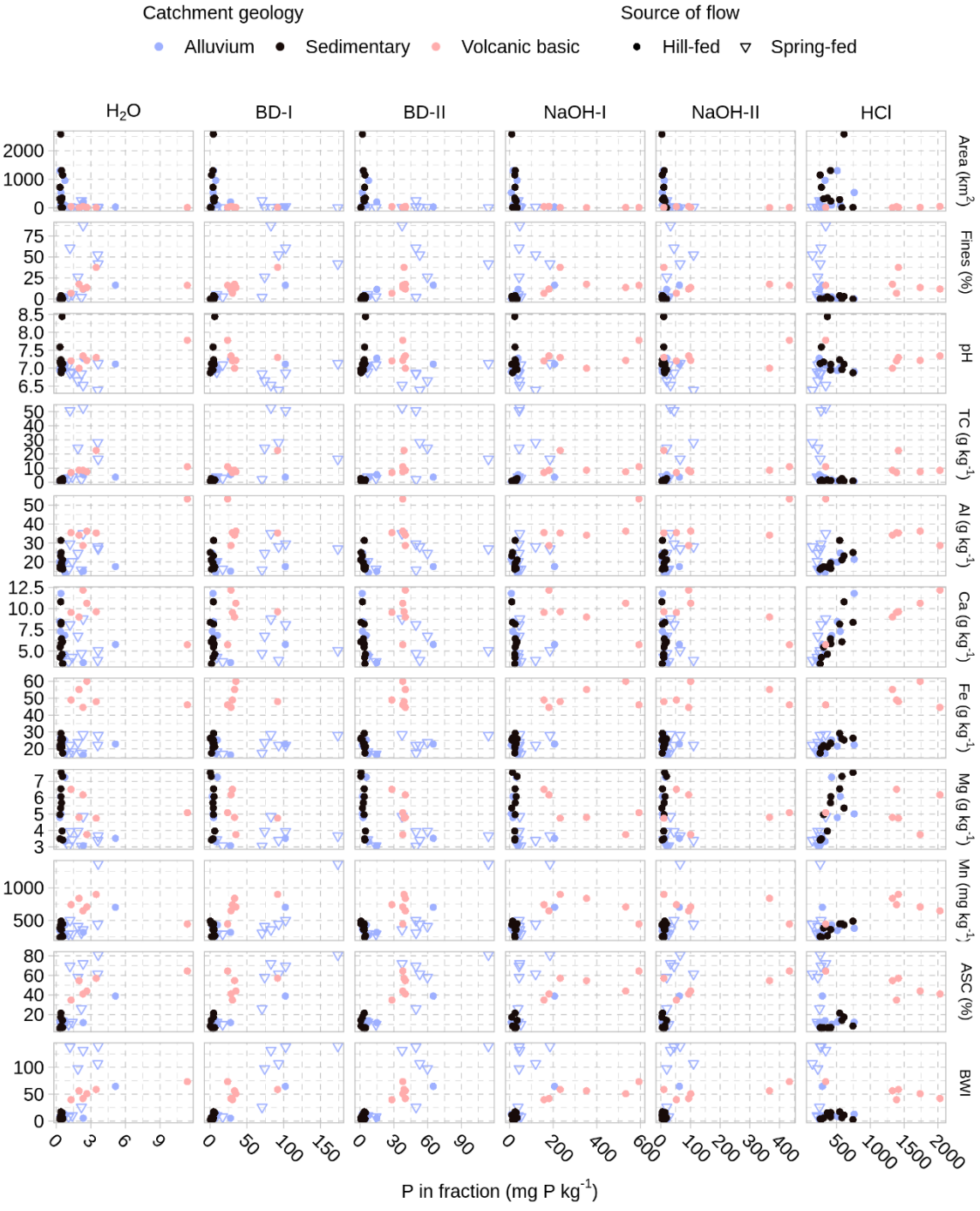


Figure S2. Log-activity of HPO_4^{2-} as a function of select ion log-activities as modelled by PHREEQC.



100

Figure S3. 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.

105

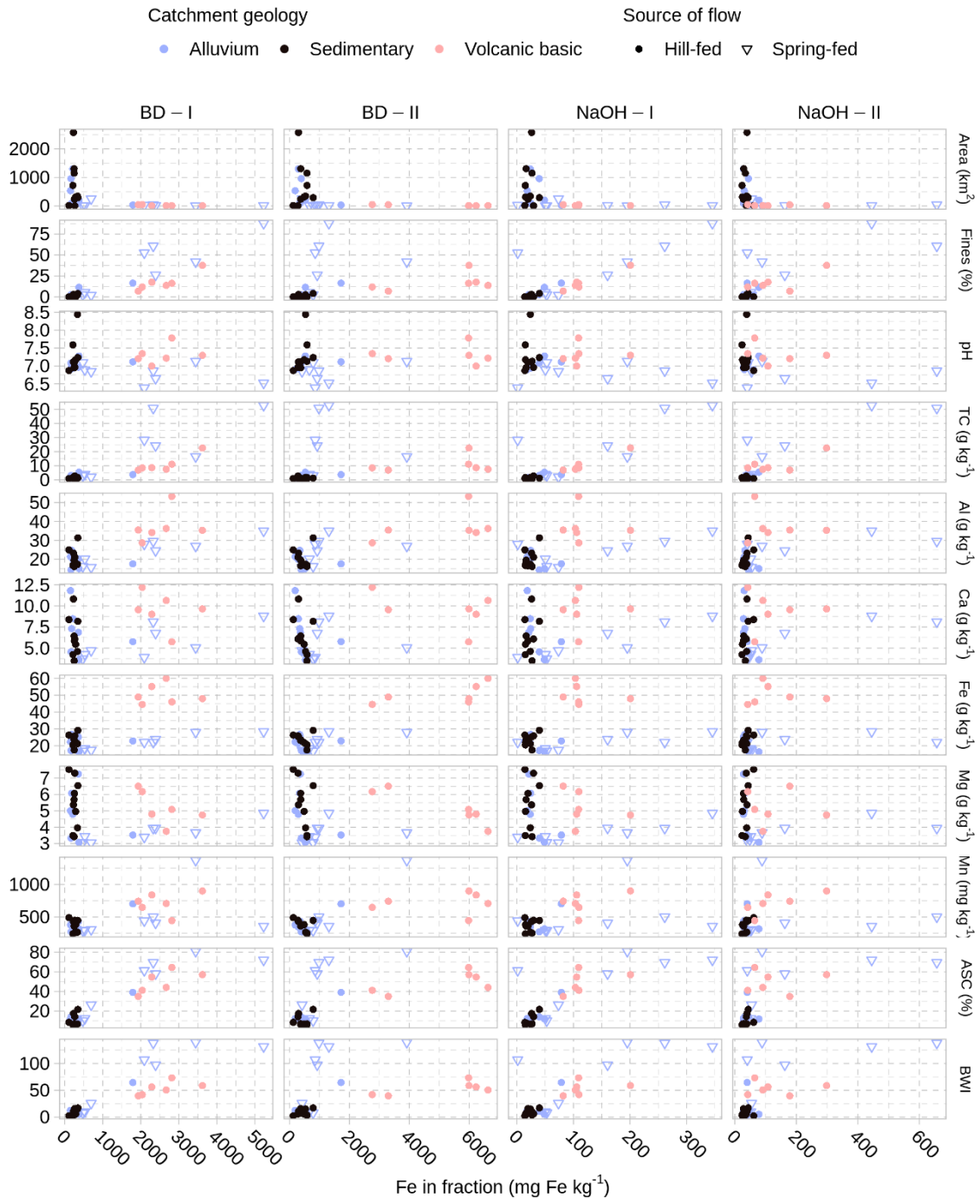


Figure S4. Scatter plots of Fe in sediment P fractions and select sediment physicochemical variables; while Fe fractions are in mg Fe kg^{-1}, each other variable has units given in the label (except Bache-Williams Index (BWI), for which we refer the reader to the main

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.

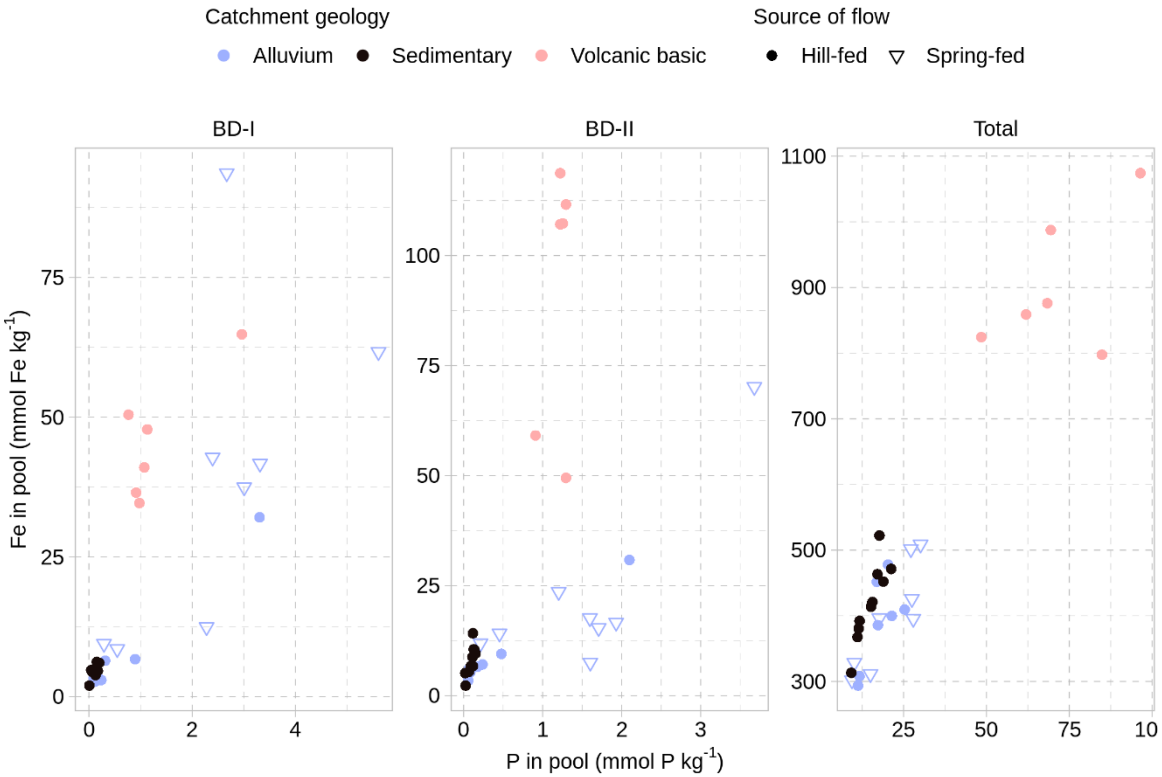


Figure S5. Sediment Fe and P content for the bicarbonate-dithionite (BD) extractions and total content; units are on molar basis to facilitate comparison of Fe:P ratios.

115

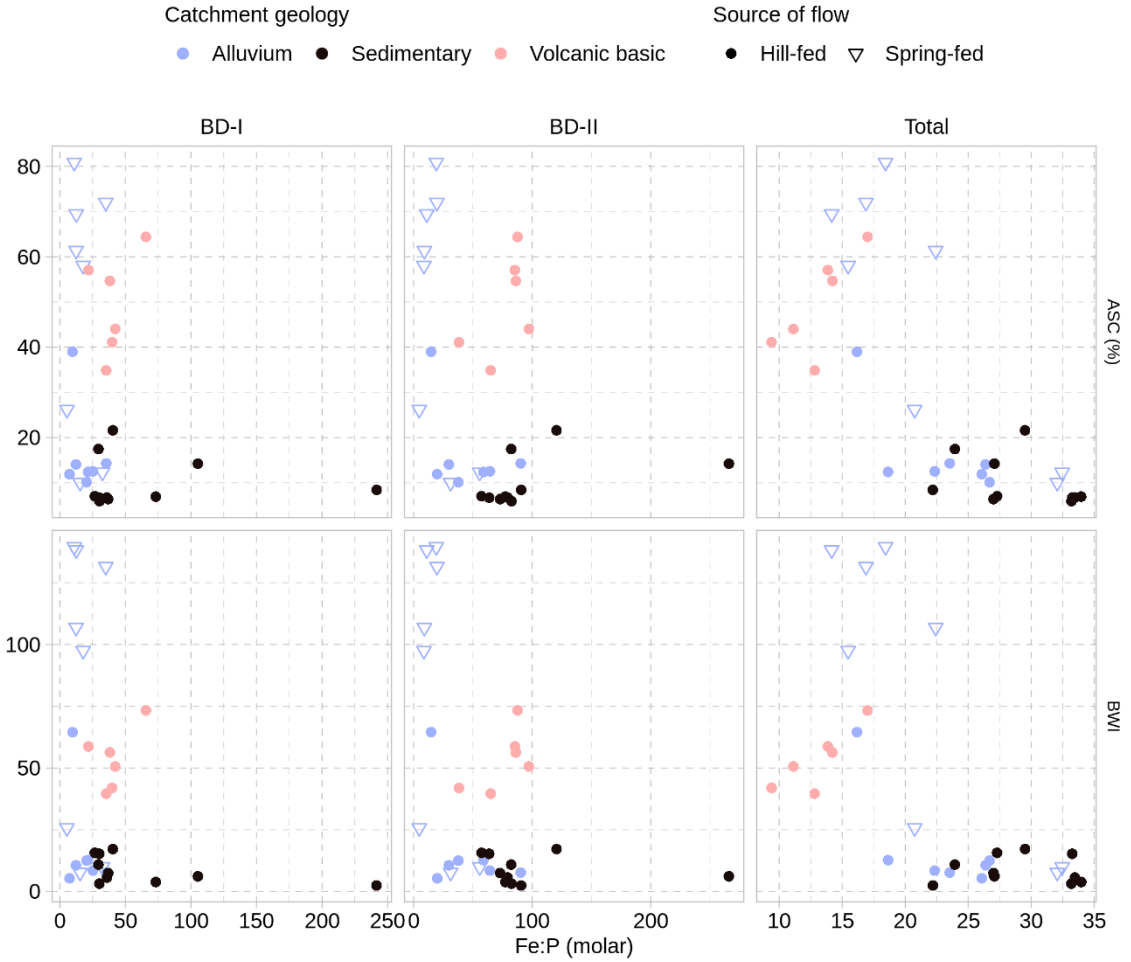


Figure S6. Sediment sorption metrics (anion storage capacity (ASC), expressed as a %, and Bache-Williams Index (BWI), whose units are given in the main text) plotted against sediment molar Fe:P ratios for the bicarbonate-dithionite (BD) fractions and total content.

Table S1. 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 (largely based on topography), geology (derived from New Zealand Land Resources Inventory 'toprock' geology data), and land-cover (1997) classifications are determined through the predominant characteristics of the catchment (full criteria details given in Snelder et al. (2010)); 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

125

Table S2. 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 S3. 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
Geology, BD-II Fe	4	15.8	269.2
Total Fe	2	21.1	283.1
Geology, Total Fe	3	17.7	276.2