

RESEARCH ARTICLE

WILEY

Solute export patterns across the contiguous USA

Dustin W. Kincaid^{1,2} | Kristen L. Underwood¹ | Scott D. Hamshaw^{1,3} | Li Li⁴ |
 Erin C. Seybold^{5,6} | Bryn Stewart⁴ | Donna M. Rizzo^{1,2,7} | Ijaz Ul Haq⁷ |
 Julia N. Perdrial^{2,8}

¹Department of Civil & Environmental Engineering, University of Vermont, Burlington, Vermont, USA

²Gund Institute for Environment, University of Vermont, Burlington, Vermont, USA

³U.S. Geological Survey, Reston, Virginia, USA

⁴Department of Civil & Environmental Engineering, The Pennsylvania State University, University Park, Pennsylvania, USA

⁵Kansas Geological Survey, University of Kansas, Lawrence, Kansas, USA

⁶Department of Geology, University of Kansas, Lawrence, Kansas, USA

⁷Department of Computer Science, University of Vermont, Burlington, Vermont, USA

⁸Department of Geography and Geoscience, University of Vermont, Burlington, Vermont, USA

Correspondence

Dustin W. Kincaid, United States Geological Survey, Upper Midwest Water Science Center, Madison, WI, USA.

Email: dustinkincaid@gmail.com

Funding information

The National Science Foundation, Grant/Award Numbers: NSF EAR-2012123, NSF EAR-2012080, NSF OIA-2033995

Abstract

Understanding controls on solute export to streams is challenging because heterogeneous catchments can respond uniquely to drivers of environmental change. To understand general solute export patterns, we used a large-scale inductive approach to evaluate concentration–discharge (C–Q) metrics across catchments spanning a broad range of catchment attributes and hydroclimatic drivers. We leveraged paired C–Q data for 11 solutes from CAMELS-Chem, a database built upon an existing dataset of catchment and hydroclimatic attributes from relatively undisturbed catchments across the contiguous USA. Because C–Q relationships with Q thresholds reflect a shift in solute export dynamics and are poorly characterized across solutes and diverse catchments, we analysed C–Q relationships using Bayesian segmented regression to quantify Q thresholds in the C–Q relationship. Threshold responses were rare, representing only 12% of C–Q relationships, 56% of which occurred for solutes predominantly sourced from bedrock. Further, solutes were dominated by one or two C–Q patterns that reflected vertical solute–source distributions. Specifically, solutes predominantly sourced from bedrock had diluting C–Q responses in 43%–70% of catchments, and solutes predominantly sourced from soils had more enrichment responses in 35%–51% of catchments. We also linked C–Q relationships to catchment and hydroclimatic attributes to understand controls on export patterns. The relationships were generally weak despite the diversity of solutes and attribute types considered. However, catchment and hydroclimatic attributes in the central USA typically drove the most divergent export behaviour for solutes. Further, we illustrate how our inductive approach generated new hypotheses that can be tested at discrete, representative catchments using deductive approaches to better understand the processes underlying solute export patterns. Finally, given these long-term C–Q relationships are from minimally disturbed catchments, our findings can be used as benchmarks for change in more disturbed catchments.

KEYWORDS

Bayesian regression, CAMELS-Chem, catchment, concentration–discharge, CONUS, export, inductive, solute

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Hydrological Processes* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Long-term observations of stream-water chemistry have been used to infer the influence of catchment structure and function on solute export by examining how solute concentrations (C) respond to changes in discharge (Q; Chorover et al., 2017). These C–Q relationships integrate signals from solute source areas within a catchment, mobilization processes that transport solutes, and transformation and retention processes occurring along flowpaths of the stream network (Gall et al., 2013; Musolff et al., 2015). Inferences derived from C–Q patterns about the influences of catchment characteristics (e.g., lithology and land cover/land use) and climate on solute loading are used to evaluate cross-catchment differences and to inform catchment management (Herndon et al., 2015; Moatar et al., 2017; Underwood et al., 2017).

Low-frequency (weekly to annual-scale) paired C–Q data are commonly used to characterize archetypes of catchment solute generation and transport. Because solute concentrations are strongly coupled to hydrology, C–Q relationships often take the power-law form and plot linearly in log–log space (Speir et al., 2024; Wymore et al., 2023). Metrics from this linear model, including the slope of the $\log(C)$ – $\log(Q)$ regression have traditionally been used to classify C–Q patterns into three main archetypes: (1) enrichment, where concentration increases with increasing discharge; (2) dilution, where concentration decreases with increasing discharge; and (3) constant (often referred to as chemostatic), where concentration is largely stable with changes in discharge (Creed et al., 2015; Godsey et al., 2009; Thompson et al., 2011).

Various unifying frameworks have been posited to explain the emergence of these C–Q archetypes, which can be generalized as the variable interaction between concentration source areas and discharge-producing zones (Basu et al., 2010; Musolff et al., 2015) across vertical (Hornberger et al., 1994; Seibert et al., 2009), lateral (Herndon et al., 2015) or longitudinal (Dupas et al., 2017) dimensions. Focus in recent years has been on vertical contrasts in subsurface solute concentration where these solutes have different origins (Zhi et al., 2019). Specifically, solutes with greater concentration at depth (e.g., geogenic solutes predominantly originating from bedrock) tend to exhibit dilution patterns in receiving streams due to the dominance of groundwater flow paths at low-flow conditions (Godsey et al., 2009; Johnson et al., 1969; Neal et al., 1990; Seibert et al., 2009; Stewart, Shanley, et al., 2022; Zhi et al., 2019). In contrast, biogenic solutes more concentrated in litter and upper soil layers tend to exhibit enrichment patterns in receiving streams due to the flushing of shallow soils at high-flow conditions (Ebeling et al., 2021; Hornberger et al., 1994; Zhi & Li, 2020). Popularized as the shallow and deep hypothesis (Zhi et al., 2019; Zhi & Li, 2020), this theory was later corroborated for a variety of solutes using process-based models (Zhi et al., 2019; Botter et al., 2020) or direct measurements of soil water and groundwater (Stewart, Shanley, et al., 2022; Zhi et al., 2019), but at one to three catchments only.

Large-sample studies of C–Q dynamics have been published in recent decades in search of an understanding of solute export

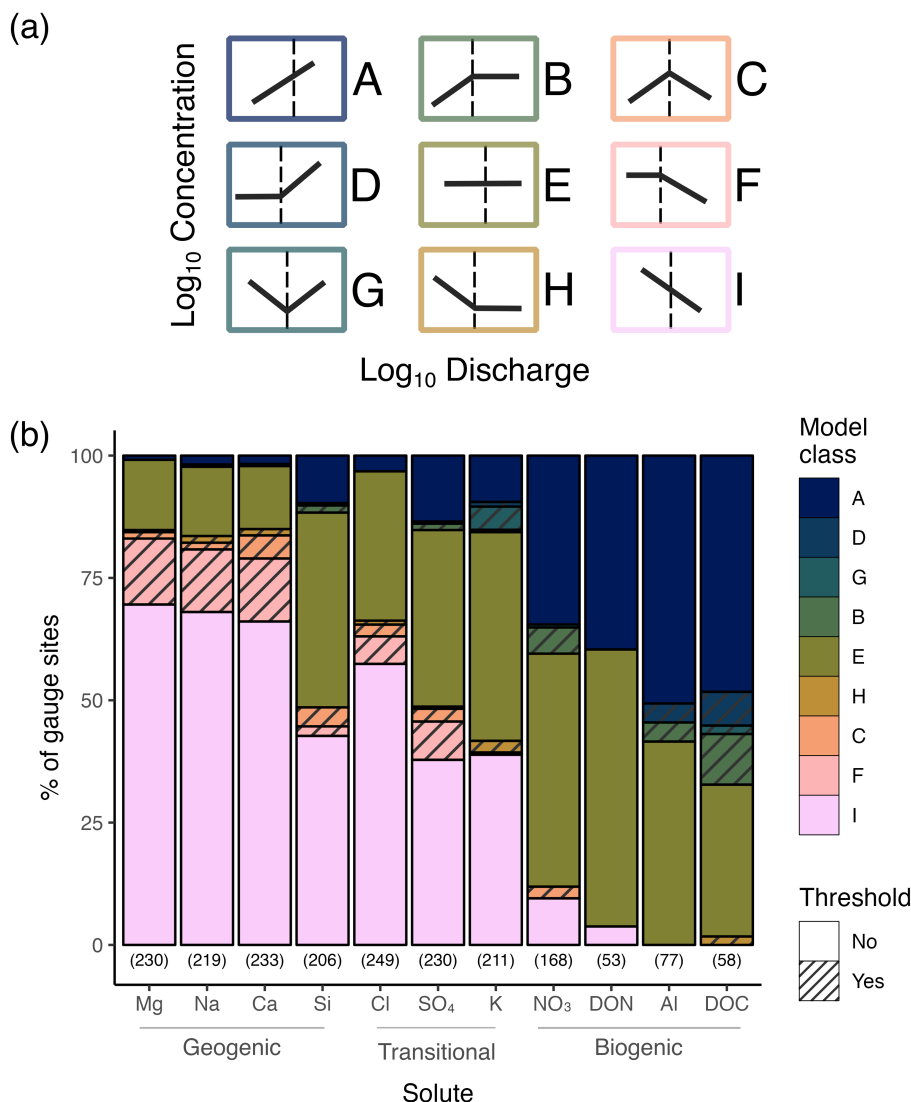
controls that is generalizable across gradients of climate, vegetation and geological characteristics. Such studies are important for overcoming the ‘uniqueness of place’ phenomenon (Beven, 2000) that results from a reliance on single site observations of stream-water chemistry to develop a generalized theory (Levin, 1992; NSF, 2018). When used to synthesize data at broad scales, such data-driven inductive approaches can identify emergent patterns that imply processes. From these process-guided interpretations of patterns, we can then generate hypotheses and eventually theories for subsequent testing at discrete, representative catchments (Underwood et al., 2023; Ward et al., 2022).

Support for the shallow and deep hypothesis as an overarching framework has been generated by large-sample studies but is generally limited to singular solutes. For example, Zhi and Li (2020) relied on reactive transport model simulations and observations from 228 CONUS catchments (7–29 000 km²) to examine nitrate C–Q patterns across gradients of climate, lithology and land use, their study demonstrated that enrichment C–Q patterns dominated in more agricultural catchments due to the relative nitrate enrichment of shallow waters as compared with deep waters. Similarly, for 278 catchments (4.4–23 200 km²) across Germany, Ebeling et al. (2021) examined C–Q patterns for nitrate, phosphate and total organic carbon, and concluded that nitrate C–Q patterns were controlled by the vertical distribution of source pools.

The shallow and deep hypothesis is appropriate for describing enrichment or dilution C–Q patterns that are linear (in log–log space) and emerge in settings where two predominant end-member source waters with relatively distinct and vertically distributed concentration sources mix (Stewart, Shanley, et al., 2022). Yet, similar C–Q patterns (equifinality of archetypes) can also emerge from laterally distributed source regions (Dupas et al., 2017; Musolff et al., 2017). This theory may not explain more complex mixing from more than two end-member source waters (Zhi & Li, 2020) or threshold patterns of C–Q responses that reflect a spatio-temporal shift in solute sourcing, transformation or mobilization (Musolff et al., 2015; Underwood et al., 2017).

The three archetypes (i.e., enrichment, dilution and constant) expand to nine modalities when more nuanced threshold dynamics are considered (Meybeck & Moatar, 2012; Moatar et al., 2017), in which the C–Q slope changes at a certain discharge to produce a segmented C–Q pattern in log–log space (Figure 1a). Initially, for ease of comparison across solutes and across catchments, the search for C–Q threshold responses was constrained by fixing the changepoint of the segmented regression at the median flow (Meybeck & Moatar, 2012; Moatar et al., 2017). A fixed discharge threshold position may adequately capture C–Q patterns for more strongly chemodynamic constituents such as nutrients and sediments (Musolff et al., 2015). However, this approach may have misestimated the frequency of threshold responses across solutes and across catchments (Moatar et al., 2017), potentially hindering a generalized understanding of the occurrence and frequency of these nonlinear solute dynamics. This fixed-threshold approach may fail to capture more nuanced threshold patterns for less chemodynamic solutes including geogenics, in

FIGURE 1 (a) C–Q archetypes A–I are segmented concentration (C)–discharge (Q) relationships proposed by Moatar et al. (2017) and adapted for data-driven threshold positions (vertical dashed lines) by Underwood et al. (2017). (b) Percentage of study catchments (gauge sites) associated with C–Q archetypes for each solute. Segmented C–Q relationships with thresholds are filled with diagonal lines. Dilution C–Q archetype I dominates geogenic responses on the left and enrichment C–Q archetype A dominates the responses for biogenic solutes on the right side. Numbers in parentheses are the total number of catchments where C–Q models were fit.



particular (Musolff et al., 2017). Previous studies have concluded that geogenic solutes are less likely to exhibit threshold C–Q patterns (Godsey et al., 2009; Musolff et al., 2015). However, this may be because the potential for threshold response has not been examined in a rigorous way using data-driven approaches across solutes, across ecoregions and for a broad range of drainage areas.

More recent studies have searched for threshold responses in a more robust, data-driven way, revealing the presence of discharge threshold values that ranged considerably above and below the median discharge. Various data-driven approaches have been used to define the discharge threshold, followed by statistical tests to confirm selection of the segmented regression over a simple regression. Previous approaches included a two-part, piecewise linear regression (Ebeling et al., 2021; Marinos et al., 2020), an iterative breakpoint(s) estimator (Diamond & Cohen, 2018; Rose et al., 2018), and Bayesian segmented regression (Underwood et al., 2017). Several of these studies have demonstrated a considerable range in threshold Q below and above the median Q, for example, for sediment and phosphorus in 18 tributaries of the Lake Champlain Basin, USA (Underwood

et al., 2017) for nitrate at 33 sites in the Mississippi River Basin (Marinos et al., 2020), and for a range of geogenic to biogenic solutes at many of 44 sites in Florida (Diamond & Cohen, 2018). Still, these studies were limited in their number of sites or geographic regions, or both.

To the best of our knowledge, only two large-sample studies (>100 sites) have included explicit examination for threshold C–Q responses at sites in Europe. Ebeling et al. (2021) evaluated both segmented and simple linear regression models for biogenic solutes (nitrate, phosphate and total organic carbon) at 278 sites (4.4–23 200 km²) in Germany; however, they chose to report and further analyse metrics from the parsimonious simple model because the segmented model performed similarly (based on Akaike Information Criterion) and generated only modest improvements in R^2 values (<10%) at up to 25% of the stations (by solute). Moatar et al. (2017) examined a range of geogenic to biogenic solutes and total suspended solids at 293 sites (50–110 000 km²) in France. Despite constraining their analysis with a fixed threshold at the median discharge, threshold C–Q patterns were detected across solute types, including up to 27% of

sites for geogenics (calcium, magnesium, silica and specific conductance).

For these long-term (annual to decadal) and low-frequency data in minimally disturbed catchments, climate and the hydrologic regime are important drivers of long-term mean concentration and mean discharge relationships across diverse catchments in the contiguous USA (i.e., arid catchments tend to have higher concentrations than humid catchments; Li et al., 2022; Stewart, Zhi, et al., 2022). However, solute C–Q relationships also depend on how catchment structure (e.g., geology, soils, topography, land cover/land use, etc.) and hydroclimatic drivers influence solute availability and the spatial heterogeneity of solute source areas and discharge-generating zones (Basu et al., 2010; Musolff et al., 2015). Still, establishing links between catchment and hydroclimatic attributes and C–Q metrics across diverse catchments can be challenging, and the strength of the links often depends on the solute characteristics (Diamond & Cohen, 2018; Ebeling et al., 2021; Pohle et al., 2021; Wymore et al., 2017). Most studies attempting to link catchment attributes to C–Q relationships have considered relatively small geographic regions that either limit hydroclimatic and/or catchment structural diversity (Ali et al., 2017; Diamond & Cohen, 2018; Ebeling et al., 2021; Musolff et al., 2015; Wymore et al., 2021). Though Godsey et al. (2019) considered these links across a global gradient of climate and lithology, they primarily focused on how arid versus humid catchments controlled the slope of the $\log(C)$ – $\log(Q)$ relationship. Godsey et al. (2019) also did not quantify C–Q threshold responses. Thus, to what extent catchment structure and hydroclimatic drivers shape segmented C–Q relationships has not been extensively examined across solutes and across the contiguous USA.

We were motivated to expand the application of segmented regression methods to a wider group of solutes across the geogenic to biogenic continuum for a greater number of catchments spanning a broader range of catchment sizes and properties across the contiguous USA. To do this, we leveraged data from the catchment attributes and meteorology for large-sample studies dataset (CAMELS; Addor et al., 2017) that was recently augmented with chemistry and paired instantaneous discharge observations for multiple solutes (CAMELS-Chem; Sterle et al., 2024). The relatively undisturbed catchments in the dataset can be considered reference sites for comparison with catchments with greater human influences (e.g., agriculture and urban sites). We used data-driven Bayesian segmented regression to identify thresholds and quantify nine C–Q archetypes (e.g., enrichment, dilution, constant or segmented patterns; Underwood et al., 2017). We also explored links between C–Q archetypes or slopes and 53 catchment attributes from CAMELS. Our overarching approach centred around inductive observation of patterns of C–Q archetypes and their metrics, including the presence of thresholds, C–Q slope b and the ratio of the coefficient of variation (CV) of concentration versus the CV of discharge across the contiguous USA. From these observations, we evaluate the following: (1) What are the discharge thresholds that trigger different behaviours for different solute groups? (2) Are there dominant C–Q archetypes and patterns of C–Q metrics for different solute groups? (3) Do hydroclimatic drivers and catchment structure relate to C–Q responses?

We used this inductive approach to highlight a few intriguing hypotheses for future deductive investigations of hydroclimatic and catchment controls on solute export.

2 | DATA AND METHODS

A schematic overview of our methods is found in Figure S1. Further, the version of the CAMELS-Chem dataset analysed here, accompanying R scripts, and Supporting Information files are available via HydroShare (Kincaid & Underwood, 2024; <https://doi.org/10.4211/hs.eddb06e91a914618a89a63bb2c2774e0>).

2.1 | CAMELS-Chem data

Our study used the recently released CAMELS-Chem database (Sterle et al., 2024). The original CAMELS dataset (Addor et al., 2017) compiled catchment attributes and meteorology data for 671 catchments from the U.S. Geological Survey (USGS) National Water Information System (NWIS; U.S. Geological Survey, 2023). These are minimally disturbed catchments filtered from the Hydro-Climatic Data Network (Lins, 2012) and represent a full range of ecoregions (Omernik, 1987). The CAMELS-Chem relational database augmented the CAMELS dataset with USGS NWIS water chemistry data for 18 common water quality constituents and instantaneous discharge from 589 of the 671 catchments and includes paired C–Q data from 1924 through 2020. We focused on 11 of the solutes and categorized them loosely based on their origins into **geogenic** (from chemical weathering; calcium [Ca], magnesium [Mg], silica [Si] and sodium [Na]), **biogenic** (from biogeochemical reactions and associations with soils; dissolved organic C and N [DOC, DON], nitrate [NO₃] and aluminium [Al]) and **transitional** groups (having both geogenic and biogenic or external origins; chloride [Cl], potassium [K] and sulphate [SO₄]).

To avoid anomalous data and C–Q responses, we restricted our analyses to USGS gaging stations where a minimum of 20 concurrent concentration (mg L^{-1}) and discharge (mm d^{-1}) measurements were collected over >3 years and spanned more than 50% of the observed range in discharge (maximum Q–minimum Q). Of the 589 catchments in CAMELS-Chem, 276 catchments met our criteria for at least one solute. However, the number of catchments varied by solute, ranging from 53 for DON to 249 for Cl (Table S1). Median record lengths across all catchments ranged from 14 years for Al to 31 years for Cl and SO₄, with a maximum record length of 95 years for Ca, Cl, Mg, Si and SO₄ (Table S1). Summaries of concentrations and discharge for each solute are also available in Table S1.

2.2 | C–Q linear and segmented regressions

To answer research questions 1 and 2, we modelled C–Q relationships for each catchment–solute combination using a data-driven Bayesian approach to simple and segmented linear regression. In addition to

the power-law function ($C = aQ^b$), we also investigated the potential presence of a single Q threshold (breakpoint in the regression) indicative of functional changes in linear C–Q relationships (Underwood et al., 2017). To estimate the power-law intercept (a), breakpoint (where present) and slope (b) of the pre- and post-threshold values, we compared simple linear and segmented regressions of the $\log_{10}(C) - \log_{10}(Q)$ relationship using a Bayesian computational approach via the ‘mcp’ package (Lindeløv, 2020; v. 0.3.1) in R (R Core Team, 2021; v. 4.1.1). For the segmented regression, we assumed that there was a single, if any, breakpoint in the C–Q relationship; and thus, we specified our segmented model with two segment formulas where the second segment had a joined slope and no change in the intercept. We evaluated model performance to determine whether the segmented regression with a breakpoint outperformed the simple regression. To do this, we compared leave-one-out cross-validation (LOOCV) and Widely Applicable Information Criterion (WAIC) values for both models (Vehtari et al., 2017) using the ‘loo’ R package (Vehtari et al., 2020; v. 2.4.1). The 95% credible intervals were estimated as the highest posterior density interval using the ‘tidybayes’ R package (Kay, 2021; v. 3.0.1) around the point estimates for intercept, breakpoint, and pre- and post-threshold slopes. Vague priors were established for all parameters so that the posterior distributions would be influenced most by the observed data (Gelman et al., 2004).

2.3 | C–Q archetype classification and metrics

To further answer research question 2, we evaluated the posterior distributions on C–Q model parameters from Section 2.2 to classify C–Q relationships for each solute into one of nine C–Q archetypes (Moatar et al., 2017; Figure 1a) following the criteria presented in Underwood et al. (2017). Positive b values for which 95% credible intervals did not span zero indicated enrichment responses. Negative b values indicated dilution responses. When b did not differ from zero (95% credible interval spans a zero value), we referred to this response as constant. Though others have referred to b values near zero as chemostatic (e.g., Creed et al., 2015; Godsey et al., 2009), we instead followed the convention of Thompson et al. (2011) and reserved the term chemostasis to describe relatively low variability in concentration as compared with discharge, consistent with others (Dupas et al., 2017; Ebeling et al., 2021; Minaudo et al., 2019; Musolff et al., 2017; Underwood et al., 2017).

We computed the ratio of the coefficient of variation (CV) of concentration versus the CV of discharge to assess the relative variability of concentration versus discharge and to quantify chemostatic versus chemodynamic behaviour (Musolff et al., 2015). Chemostatic behaviour occurs when discharge variability is high relative to concentration variability and often arises from a homogenization of solute stores in the catchment (Basu et al., 2010; Musolff et al., 2017; Thompson et al., 2011). Chemodynamic behaviour occurs when concentration variability is high relative to discharge variability and may arise from temporally variable connectivity between solute sources and streamflow generating zones in the catchment (Basu et al., 2010; Bende-Michl et al., 2013).

2.4 | Spatial distributions of C–Q relationships and links to catchment attributes

2.4.1 | Hierarchical clustering of catchment attributes

To understand how C–Q relationships varied among catchments with similar attributes (i.e., research question 3), we first clustered catchments into groups using 53 numerical and categorical catchment attributes from the original CAMELS dataset that likely link to C–Q relationships (Table S2). We also calculated the mean surface water temperature and pH from the chemistry dataset (Section 2.1). We clustered catchments using an agglomerative hierarchical clustering analysis. Prior to clustering, we imputed the few missing continuous values (~1% of all values) using the expectation-maximization algorithm on five bootstrapped samples of the incomplete data with the ‘Amelia’ R package (Honaker et al., 2011; v. 1.8.0). We then created the dissimilarity matrix for the mixed variable types using the ‘daisy’ function in the ‘cluster’ package (Maechler et al., 2021; v. 2.1.2) in R using Gower’s (1971) general dissimilarity coefficient. Here, the dissimilarity between two rows is based on the weighted mean of the contributions of each variable. We performed the clustering analysis using Ward’s (1963) clustering criterion via the ‘hclust’ function in base R. Finally, we selected the optimal number of clusters using the elbow curve method (Thorndike, 1953).

2.4.2 | Statistics for differences in C–Q relationships across clusters

Continuing with our approach for research question 3, we conducted chi-square tests of independence to examine differences in the distributions of C–Q archetypes across the catchment clusters for each solute. To do this, we used the ‘vcd’ R package (Meyer et al., 2021; v. 1.4–9), which conducts the test of independence and visualizes the results as a mosaic plot (graphical representation of a contingency table showing proportions of data in each group) with Pearson residual-based shadings and labels for visualizing conditional independence (Zeileis et al., 2007). Residual values greater than 2 or less than -2 represent a significant departure from independence.

To test for differences in the C–Q relationship slope b and CV ratios (only considering C–Q archetypes without thresholds, i.e., A, E and I) across catchment clusters, we conducted non-parametric Kruskal–Wallis rank sum tests. Following the rejection of the null hypothesis ($\alpha = 0.05$), we conducted pairwise comparisons of b and CV ratios between clusters using two-sided Conover–Iman tests ($\alpha = 0.05/2$). To control the false discovery rate, we adjusted p -values using the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995).

2.4.3 | Quantifying links between catchment attributes and C–Q relationships

We evaluated the strength of the relationships between catchment attributes and both C–Q archetype and C–Q slope b using random

forests for feature importance. We also calculated Spearman's rank correlation coefficients between catchment attributes and the C–Q slope b for C–Q archetype classifications without thresholds (i.e., A, E and I) for each solute. However, there were no obvious patterns within solute groups (i.e., geogenic, transitional, and biogenic) in terms of common important attributes or categories of attributes (e.g., climate, hydrology, land cover, etc.) for predicting C–Q archetypes or slope b . As such, we do not discuss these results extensively, and all detailed methods and results have been moved to Supporting Information.

3 | RESULTS

3.1 | C–Q responses with thresholds were relatively rare

C–Q archetypes were dominated by responses with no thresholds (C–Q archetypes A, E and I; Figure 1), representing 79%–100% of solute-specific archetypes. We also considered two alternate segmented C–Q archetypes for both models A and I where post-threshold slopes were non-zero (not constant; Figure S3); however, these alternate segmented archetypes comprised only 2% of all C–Q responses. Thus, we aggregated these alternate segmented archetypes into the corresponding A and I C–Q archetypes, and do not discuss these further.

Accordingly, segmented C–Q archetypes (archetypes i.e., B, C, D, F, G and H) were relatively rare, comprising only 8% of C–Q responses for Al, NO₃ and Si, and a maximum of 19% and 21% for Ca and DOC, respectively. Geogenic solutes (Ca, Mg, Na and Si) accounted for the majority (56%) of threshold responses. For three of the four geogenic solutes, Ca, Mg and Na, and two transitional solutes, Cl and SO₄, constant-dilution C–Q archetype F occurred more frequently than other segmented responses (6%–13% of models for each solute). For the geogenic solute, Si, the most common segmented C–Q archetype was enrichment-dilution C–Q model C (4% of the catchments). For the transitional solute, K, dilution-enrichment segmented C–Q archetype G occurred most frequently (5% of catchments). For three biogenic solutes, Al, DOC and NO₃, enrichment-constant segmented C–Q archetype B occurred more frequently than other segmented responses (4%–10% of all C–Q archetypes for each solute).

Flow-exceedance probabilities (i.e., 100 – streamflow percentile) associated with threshold onset in segmented C–Q archetypes ranged from 2% to 98% with a median flow-exceedance probability of 32% (i.e., above average streamflow) across all solutes (Figure S4; Table S3). In other words, most solutes were dominated by thresholds that occurred at flows greater than median flow. Specifically, 57% to 100% of threshold responses depending on solute occurred at flows greater than median flow. The only exception was SO₄, which had one threshold that occurred above and another below median flow. None of the catchments had more than one solute with a threshold C–Q response.

3.2 | C–Q responses

3.2.1 | Geogenic solutes (Ca, Mg, Na and Si)

Geogenic solutes tended to be dominated or co-dominated by dilution C–Q responses. For Ca, Mg and Na, dilution archetype I occurred in 66%–70% of catchments (Figure 1; Tables S4 and S5). Si, on the other hand, was co-dominated by dilution archetype I and constant archetype E, which occurred in 43% and 40% of catchments, respectively. C–Q slope b values (for C–Q archetypes A, E and I only; Figure 2a,d; Table S6) corroborated these patterns, specifically that geogenic solutes were dominated by dilution C–Q responses (negative b values). However, dilution responses were relatively weak. Ca, Mg and Na had median b values of -0.13 (Ca and Mg) to -0.12 (Na) across all catchments, though the interquartile ranges (IQR; range from the first to the third quartile) remained negative. The median b for Si was even less negative (-0.03), and the IQR spanned zero (-0.11 to 0.01). In addition to their dilution C–Q responses, the geogenic solutes were exported chemostatically (CV ratio <0.7 when considering C–Q archetypes A, E and I only) in 95% to 100% of catchments depending on solute, with median CV ratios of 0.19, 0.19, 0.20 and 0.15 for Ca, Mg, Na and Si, respectively (Figure 2b,d; Table S6).

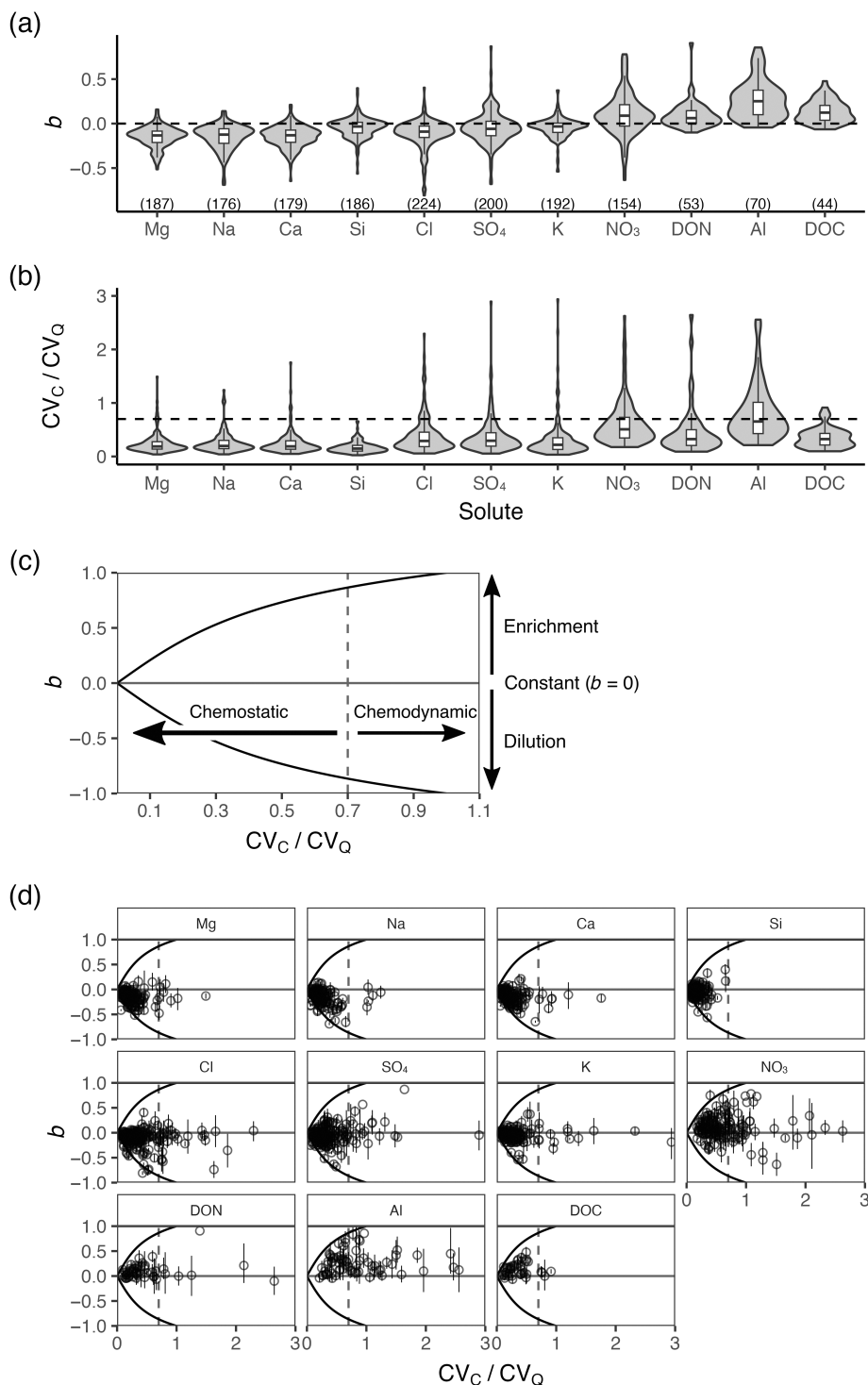
3.2.2 | Transitional solutes (Cl, K and SO₄)

C–Q responses for K and SO₄ were co-dominated by constant C–Q archetype E and dilution archetype I, whereas Cl was dominated by dilution archetype I in 57% of catchments (Figure 1; Tables S4 and S5). Only 31% of catchments had constant archetype E responses for Cl. Correspondingly, the median C–Q slope b value for Cl (-0.09) was slightly more negative than b values for K and SO₄ (-0.03 and -0.06 , respectively; for C–Q archetypes A, E and I only; Figure 2a,d; Table S6). Reflecting the dominance of dilution archetype I for Cl, the IQR did not span zero (-0.16 to -0.03), but it did for K and SO₄ (K: -0.10 to 0.01 ; SO₄: -0.14 to 0.03). Like geogenic solutes, all three transitional solutes were exported chemostatically in 86%–93% of catchments depending on solute, with median CV ratios of 0.29, 0.22 and 0.30 for Cl, K and SO₄, respectively (Figure 2b,d; Table S6).

3.2.3 | Biogenic solutes (Al, DOC, DON and NO₃)

In contrast to geogenic and transitional solutes, biogenic solutes were dominated by C–Q archetype E or enrichment archetype A or both. Constant archetype E was dominant for DON and NO₃, occurring in 57% and 48% of catchments, respectively. The second most common response for DON and NO₃ was enrichment archetype A (40% and 35% of catchments, respectively). Al and DOC responses were most frequently enrichment archetype A (51% and 48%, respectively), though a large proportion of catchments had constant archetype E responses for these solutes (42% and 31%, respectively). Reflecting the shift to either constant or enrichment C–Q responses, the median b values were all

FIGURE 2 For concentration (C)–discharge (Q) relationship archetypes without thresholds (A, E, I; Figure 1): (a) the distributions of slope b of the C–Q relationship and number of catchments in parentheses; (b) the ratio of the coefficients of variation (CV) for C and Q. The horizontal dashed line at 0.7 represents the threshold for chemostatic (CV ratio <0.7) and chemodynamic (CV ratio >0.7) responses. The CV ratio axis was limited to values between 0 and 3 for visual clarity, and thus 4 CV ratios >3 are not shown (Cl = 3.3, 3.6; Mg = 4.4; SO₄ = 3.5). Violin plots (shaded regions) represent the mirrored density distribution of the values. Boxplots within the violin plots represent the median and interquartile range. (c) Example figure to guide interpretation of plots in (d). (d) Slope b versus the CV ratio to visualize export regimes. Vertical error bars represent 95% credible intervals for b . The bounds in the upper and lower left areas of plots were defined solely by CV_Q and b (not CV_C), and have been derived from the mean and standard deviation of Q from all data (Musolff et al., 2015). The CV ratio axis was also limited to values between 0 and 3 for visual clarity.



positive, with a minimum of 0.06 for DON and a maximum of 0.25 for Al (for C–Q archetypes A, E and I only; Figure 2a,d; Table S6). NO₃ was the only biogenic solute with an IQR that spanned zero (−0.03 to 0.21). Generally, biogenic solutes shifted towards constant and enrichment C–Q archetypes and positive b values relative to geogenic and transitional solutes. However, only two of the solutes, NO₃ and Al, had IQRs for CV ratios that spanned 0.7 (third quartile: 0.73 and 1.01, respectively, when considering C–Q archetypes A, E and I only) and were exported chemodynamically in 29% and 44% of catchments, respectively (Figure 2b,d;

Table S6). The remaining biogenic solutes, DOC and DON, both had median CV ratios of 0.33 and were exported chemodynamically in only 9% and 13% of catchments, respectively.

3.3 | Hierarchical clustering of study catchments

Clustering the catchments using CAMELS attributes (see text in Supporting Information) resulted in five distinct clusters generally

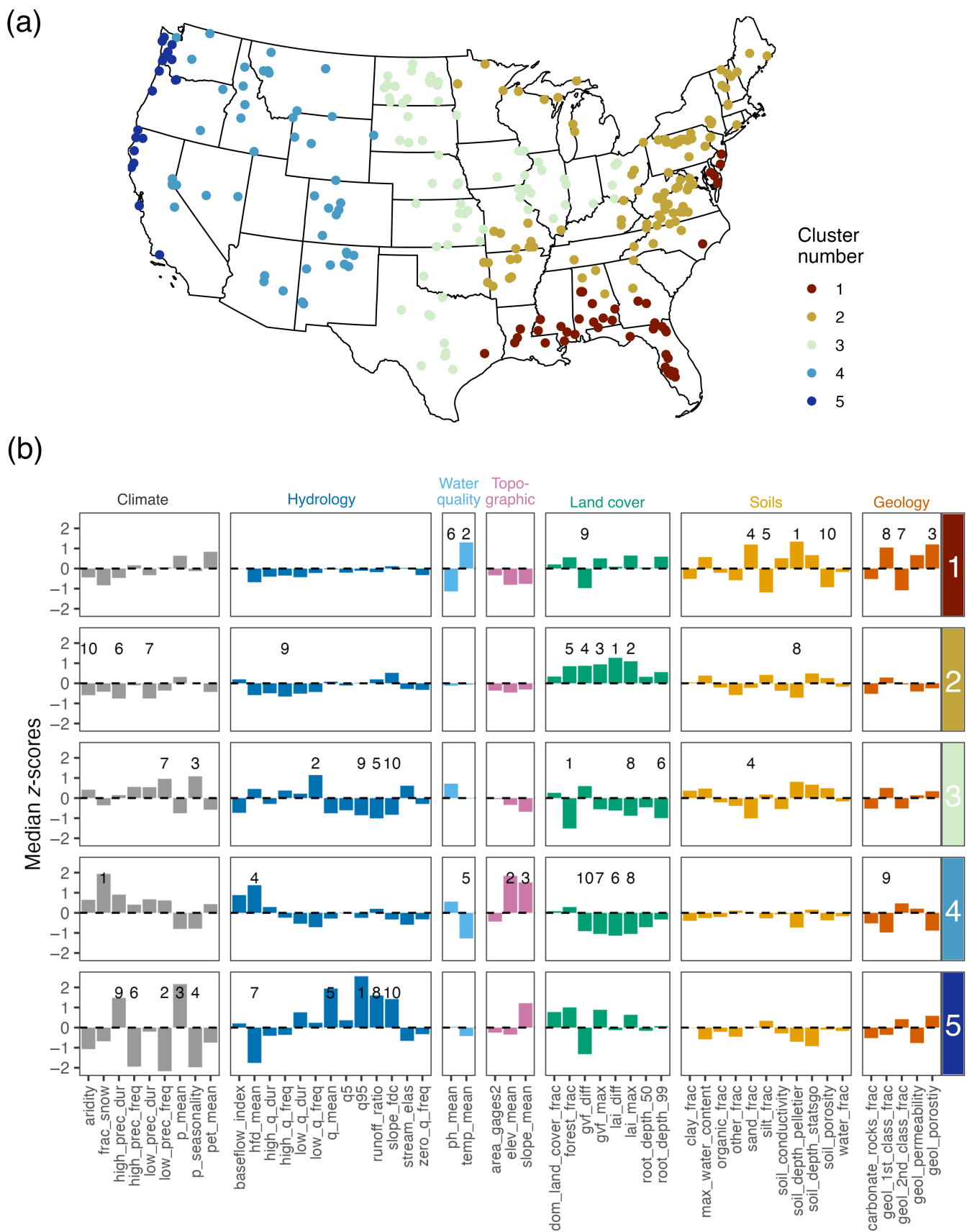


FIGURE 3 Legend on next page.

segregated by geographic location (Figure 3; see Figure S5 for elbow plot used to determine the optimal cluster number).

3.4 | Linking geographic location to C–Q responses

3.4.1 | Geogenic solutes (Ca, Mg, Na and Si)

Most departures from typical export behaviour for geogenic solutes occurred in the first three clusters located in the central and eastern USA, with the most divergent responses occurring in cluster 3 (Figures 4b and S7). Further, most departures occurred for dilution-dominated Ca and Mg. In general, cluster 3 had the highest proportion of threshold responses across all clusters (24%; 120 of 497 C–Q archetypes in cluster 3). Of the 120 threshold responses in cluster 3, 64% of these were for geogenic solutes, accounting for 33% of all geogenic C–Q archetypes in cluster 3. For example, Ca and Mg had fewer dilution archetype I responses (Pearson residuals [PR]: -2.9 and -2.4 for Ca and Mg, respectively; PR values greater than 2 or less than -2 represent a significant departure from independence), more enrichment-dilution archetype Cs (PR: 3.6 and 2.5, respectively), and constant-dilution archetype Fs (PR: 3.6 and 5.2). Mg also had more constant archetype E responses than expected (PR: 2.0). Similar to Ca and Mg, Na had more constant-dilution archetype F responses in cluster 3 (PR: 4.1). For Ca, Mg and Si, the highest median b occurred in cluster 3 (-0.07 , -0.11 and 0.00 , respectively; Tables S6 and S7). For Si, cluster 3 was the only cluster where the median b was not negative. The median CV ratio for all geogenic solutes across all clusters was <0.7 (Figure S2; Table S6).

3.4.2 | Transitional solutes (Cl, K and SO₄)

Most departures from typical export behaviour for transitional solutes also occurred in the first three clusters. Most of the departures were for SO₄, many of which reflected shifts in Ca and Mg export behaviour. For example, in cluster 1, SO₄ had a larger proportion of constant C–Q archetype E (PR: 2.6; Figures 4b and S7) and fewer dilution archetype I responses (PR: -2.7). Consequently, SO₄ had the highest (and only positive) median b in cluster 1 when considering C–Q archetypes without thresholds (i.e., A, E and I; Figure 4d; Tables S4 and S5).

In cluster 2, like Mg, SO₄ had fewer constant-dilution archetype Fs (PR: -2.4). Median b values for SO₄ were second highest in cluster 2, continuing an east-to-west trend in decreasing median b values for SO₄ where the most negative median b occurred in cluster 5 (though the median was only significantly different than the median in cluster 1; Tables S6 and S7). Like geogenic solutes, cluster 3 had the most departures from normal for transitional solutes. And, like Ca and Mg, SO₄ had more than expected enrichment-dilution archetype C (PR: 3.1) and constant-dilution archetype F responses (PR: 4.1). Consequently, SO₄ had fewer than expected constant archetype E responses (PR: -2.2). Further, as was the case for Ca, Mg and Si, K had the highest median b in cluster 3 (-0.01). Conversely, in adjacent cluster 4, the most negative median b value occurred for K, similar to Ca, Mg and Na.

3.4.3 | Biogenic solutes (Al, DOC, DON and NO₃)

In general, no meaningful departures from typical export behaviour occurred for DON. For Al, DOC and NO₃, departures occurred across all clusters except cluster 2. Most of the departures were for NO₃. For example, there were fewer enrichment C–Q archetype A responses (PR: -2.1 ; Figures 4b and S7) and more dilution archetype I behaviour (PR: 2.0) for NO₃ in cluster 1. Correspondingly, median b values for NO₃ (-0.04) were lowest in cluster 1, and it was the only cluster where NO₃ median b values were negative (Figure 4d; Tables S6 and S7). In general, there was an east-to-west trend in increasing median b values for NO₃ where the highest median b occurred in cluster 5 (0.26) and was 1.7 times greater than the next highest median b value for NO₃ (cluster 3: 0.15). Conversely, the highest median b for Al (0.63) occurred in cluster 1 and was 1.5 times greater than the next highest median slope b value of 0.41 in cluster 5. For DOC, the highest median b (0.20) occurred in cluster 4 and the lowest in the adjacent cluster 3. Median CV ratios for biogenic solutes also departed from typical values in a few clusters (Figure S2; Tables S6 and S8). Cluster 3 had the only median CV ratio <0.5 for NO₃. Conversely, cluster 5 had the only median CV ratio >0.5 for DOC, though the median was only significantly different from clusters 1 and 3 (Table S8). Cluster 5 also had the only median CV ratio >1 for Al, but there were no significant statistical differences in median CV ratios across catchment clusters.

FIGURE 3 (a) Study catchments coloured by cluster number determined from hierarchical clustering of CAMELS attributes (Table S2). (b) Median z-scores (the number of standard deviations by which the median attribute value for a cluster of catchments is above or below the mean attribute value across all clusters) for all numerical critical zone attributes. Each row of plots corresponds to a cluster of catchments in (a) and the colour of each row title panel on the right corresponds to the cluster number shown in (a). The colour of each bar in the plots corresponds to the general category of each attribute. The numbers in each plot indicate the 10 catchment attributes with the highest absolute median z-score for each cluster of catchments, where 1 is the highest and 10 the lowest. Boxplots of the raw value ranges for each attribute in each cluster are shown in Figure S6a. The distribution of catchments among levels of the categorical critical zone attributes (dom_land_cover, geol_1st_class, geol_2nd_class, high_prec_timing and low_prec_timing; Table S2) used in the hierarchical clustering of catchments is shown in Figure S6b.

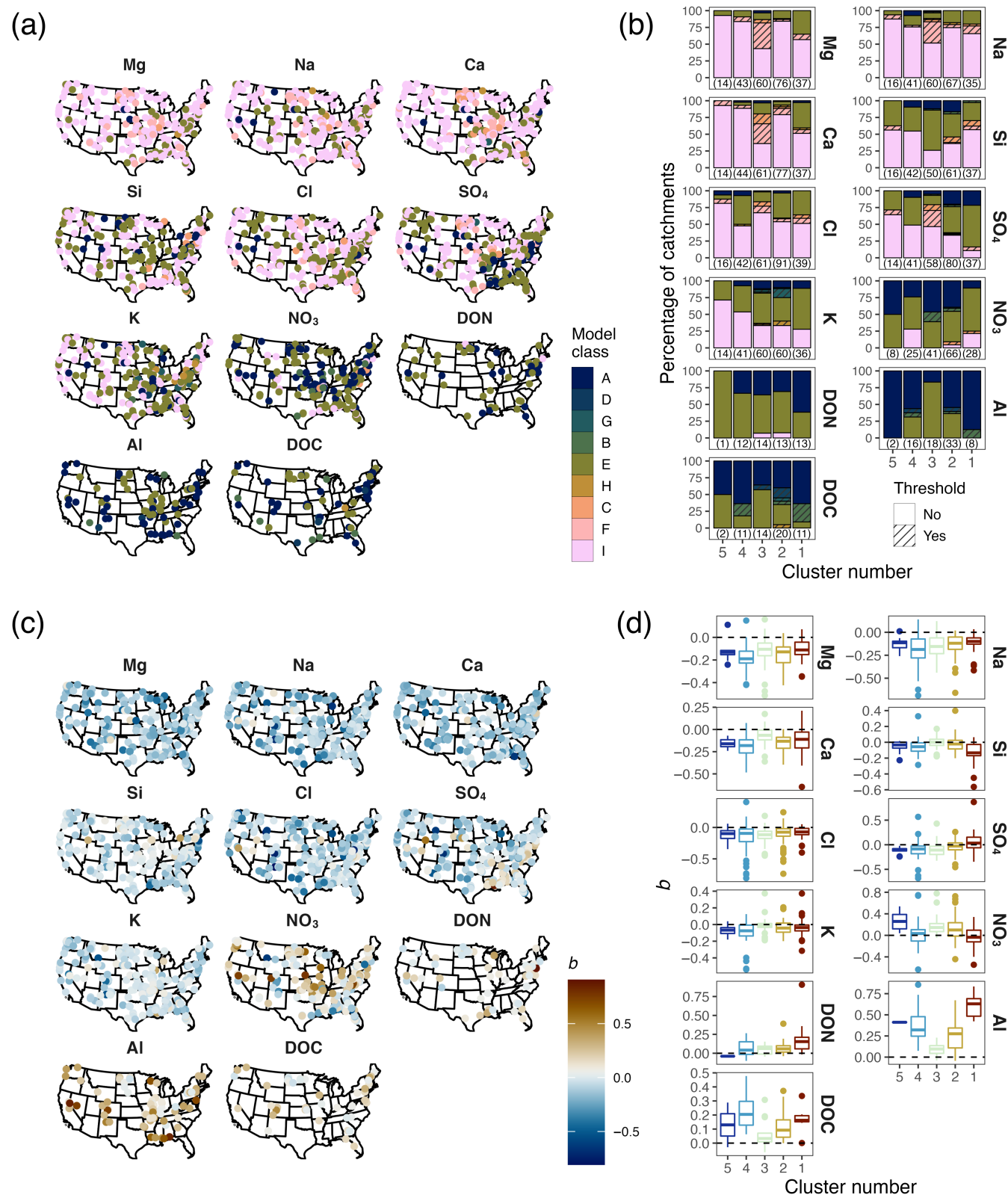


FIGURE 4 (a) Study catchments coloured by concentration (C)–discharge (Q) archetype for each solute. (b) Distributions of C–Q archetypes across clusters shown in Figure 3a. Segmented C–Q relationships with thresholds are filled with diagonal lines. Numbers in parentheses are the total number of catchments where C–Q models were fit for each cluster. (c) Study catchments coloured by the slope b of the C–Q relationship for C–Q archetypes without thresholds (A, E, I; Figure 1) for each solute. (d) Distributions of slope b across clusters in (a). Boxplots represent the median and interquartile range. Results of the chi-square tests of independence to test for differences in the distribution of C–Q archetypes across catchment clusters are visualized in Figure S7. Results of Kruskal–Wallis rank sum tests and Conover–Iman tests of multiple comparisons using rank sums to test for differences in the slope b across catchment clusters are in Table S7.

3.5 | Linking catchment attributes to C–Q responses

The most important catchment attributes for classifying C–Q archetypes were unique to each solute. In other words, there were no obvious patterns within solute groups (i.e., geogenic, transitional and biogenic) in terms of common important attributes or categories of attributes (e.g., climate, hydrology, land cover, etc.) for predicting C–Q archetypes (see Supporting Information and Figure S8). Similarly, there were no obvious patterns for predicting slope b among solute groups, as measured by random forest regression models and Spearman's rank correlations (Figure S8b and S9).

There were also no strong relationships between catchment size and the chemostatic or chemodynamic export regime of solutes in this study (Figure S10).

4 | DISCUSSION

4.1 | Threshold patterns in C–Q relationships are rare and vary in their distribution across space and solute

Threshold responses in C–Q relationships are important to understand because they provide insight into different functional stages of solute flux and can impact the accuracy of predictive models that are needed to inform management strategies (Diamond & Cohen, 2018; Pohle et al., 2021). Across catchments and solutes in this study, threshold responses were rare, representing only 12% of responses (233 of 1934) across our 276 catchments of the contiguous USA. Of these responses, 56% occurred for geogenic solutes, and only 14% for biogenic solutes. In a changing climate where increased frequency and magnitude of storms are projected, increasing concentrations at some Q threshold could present water quality challenges; however, such post-threshold flushing patterns accounted for only 1% (24 of 1934) of C–Q responses for solutes assessed in this study. When thresholds were detected, they occurred at flows greater than median discharge (Figure S4). Thus, our results contradict previous studies suggesting that most threshold responses occur at median discharge (Diamond & Cohen, 2018; Moatar et al., 2017), and are consistent with Marinos et al. (2020) who found a range of discharges for NO₃ C–Q threshold responses in the Upper Mississippi River Basin, USA.

The frequency of solute threshold responses observed in this study was on the lower end of what has been observed in western Europe where researchers explored C–Q responses for multiple solutes across a large number of catchments. In France, 46% of catchment–solute combinations had threshold responses for C–Q relationships (Moatar et al., 2017). The solute with the least number of threshold responses, Si, still had threshold responses in 21% of the observed catchments. However, Moatar et al. (2017) may have misestimated the frequency of threshold responses because they restricted their search for C–Q threshold responses by fixing the changepoint of their segmented regressions at median flows. In Germany where

researchers focused on biogenic solutes, 25% of catchment–solute C–Q responses were threshold responses (Ebeling et al., 2021).

The lower frequency of threshold responses in this study as compared to western Europe may result from differences in the distribution of land uses between the two datasets. Catchments in both western European studies were frequently dominated by urban and agricultural land uses, whereas the original CAMELS catchments were intended to represent minimally disturbed landscapes. When others have focused on catchments that are more intensively impacted by agriculture in our cluster 3 region (Figure 3a), they found that NO₃ C–Q relationships frequently exhibited threshold responses as a result of agricultural practices (Ma et al., 2022; Marinos et al., 2020). Specifically, nitrogen fertilization creates vertical nitrogen stratification in agricultural soils where baseflows are sourced from lower-nitrogen subsurface layers and event flows are primarily sourced from shallow soils with higher nitrogen content (Ma et al., 2022). Tile drainage can exacerbate the two-stage C–Q relationship by serving as a conduit for shallow soil water with uniformly high nitrogen concentrations at higher discharges (Cain et al., 2022; Ma et al., 2022; Marinos et al., 2020).

The lower frequency of threshold responses in our study may also relate to the number of paired C–Q samples for each site–solute relationship. In this study, catchments with detected threshold C–Q relationships were characterized by more observations (median: 115; IQR: 55–170) than catchments where no threshold response was identified (median: 70; IQR: 35–154). The studies in France and Germany limited their analyses to gauging stations with a minimum of 300 and 70 paired concentration and discharge observations, respectively. Conversely, the minimum number of paired observations in our study was set at 20 following Zarnetske et al. (2018). Thus, having more samples at sites that capture the full range of solute dynamics arising from variable seasonal or catchment wetness conditions (Burns et al., 2019; Knapp et al., 2022) may allow us to more robustly characterize any threshold patterns that exist.

Of the locations where we observed threshold C–Q responses, cluster 3 in the central USA (Figure 3a) had the highest proportion of threshold responses (24% of all responses in cluster 3). The majority (67%) of the threshold responses in cluster 3 were constant–dilution C–Q archetype F responses (Figure 1a), and Ca, Cl, Mg, Na and SO₄ had C–Q archetype F responses in 10%–38% of cluster 3 catchments depending on solute. In France where studied catchments were frequently dominated by urbanization and agriculture, Ca, Mg and conductivity exhibited the same threshold response in ~25% of catchments (Moatar et al., 2017). Despite targeting minimally disturbed catchments, the dominant land use in >50% of CAMELS-Chem catchments in cluster 3 (38 of 67) was a mixture of cropland and natural vegetation or just croplands (Figure 3b). Of the solute–catchment C–Q responses in cluster 3 that were constant–dilution archetype F, 70% (56 of 80) had the cropland mixture or just croplands as the dominant land use. As such, the constant–dilution threshold response may be typical of catchments dominated by agriculture. Additionally, cluster 3 catchments are also underlain by unconsolidated, carbonate-rich sediments (Figure 3b) and have a high probability of streamflow

intermittency (Messenger et al., 2021) with frequent low or no-flow states and/or groundwater-driven contributions to streamflow. Therefore, the constant-dilution threshold response may emerge from the frequent occurrence of low streamflow that is dominated by deeper groundwater enriched with base anions and cations from chemical weathering at depth. As discharge increases, streamflow becomes diluted by shallow groundwater from weathered unconsolidated sediments and soils that are less enriched with geogenic anions and cations. Tile drainage could also produce a similar effect in diluting ion-rich deeper groundwater as discharge increases. Thus, given the mixture of dominant land covers/land uses among catchments with the constant-dilution C–Q responses, agricultural practices may not be the only driver of these threshold behaviours, but they may amplify them in the landscape.

4.2 | Dominant C–Q patterns differ across solute groups

Dominant C–Q patterns differed across solute groups. Specifically, geogenic solutes Ca, Mg, Na and Si largely exhibited weakly diluting C–Q responses across the study catchments. The geogenic solutes were also generally exported chemostatically (Figure 2b,d). These results correspond to findings from previous studies that examined a similarly broad range of catchment sizes across gradients in climate, topography and geology (Godsey et al., 2009; Godsey et al., 2019; Moatar et al., 2017). Specifically, C–Q responses for geogenic solutes are weakly diluting to quasi-constant.

Transitional solutes, Cl, K and SO₄, have both geologic and biogenic sources, and C–Q responses for these solutes were co-dominated by dilution and constant responses. They were also exported chemostatically across most catchments, though chemodynamic responses occurred more frequently than for geogenic solutes, possibly because these solutes have multiple sources including both atmospheric deposition and subsurface rock dissolution. For example, Cl and SO₄ can be deposited atmospherically with rain (Berner & Berner, 2012) and originate from weathering of silicates, shale and evaporites (Diamond & Cohen, 2018; Mayer et al., 2010; Musolff et al., 2015). K is a nutrient present in fertilizers, can leach from organic matter, be deposited atmospherically and originate from silicate weathering at depth (Alfaro et al., 2004; Likens et al., 1994; Sardans & Peñuelas, 2015).

Conversely, biogenic solutes more frequently exhibited enrichment C–Q responses, reflecting lower concentrations at baseflow than at high flows as the shallow and deep hypothesis predicts. However, DON and especially NO₃ had more constant and negative C–Q responses than DOC or co-transported Al. Variability in NO₃ C–Q responses within and among catchments has been demonstrated by others (Moatar et al., 2017; e.g., Ebeling et al., 2021; Knapp et al., 2022) and reflects the heterogeneity in vertical NO₃ concentration gradients among catchments that result from the dynamic nature of NO₃ cycling and transport throughout the subsurface (Ebeling et al., 2021). However, others have shown that variability of biogenic

solute (e.g., DOC, nutrients and metals) C–Q responses can result from temporal variability in lateral hydrologic connectivity of the stream to the adjacent riparian area/floodplain mediated by catchment wetness (Boyer et al., 1997; Knapp et al., 2022; Vidon & Hill, 2004). Catchment wetness interacts with the large lateral and vertical heterogeneity in biogenic solute concentration, increasing concentration variability relative to discharge variability and driving the chemodynamic behaviour more frequently observed for solutes associated with soils (Knapp et al., 2022). Alternatively, or likely in addition to the vertical gradient of biogenic solutes in the subsurface, the pulse-shunt concept (Raymond et al., 2016) may explain the increased frequency of enrichment C–Q responses for more bio-reactive solutes. Specifically, high-discharge events reduce the residence time of bio-reactive solutes, limiting the ability for biological uptake of these solutes during larger hydrologic events.

4.3 | Data-driven, inductive analysis yields hypotheses to be tested using deductive approaches

Our focus in this study has been on a data-driven analysis to evaluate the occurrence and frequency of solute C–Q metrics and response types, including threshold responses, across broad scales of the contiguous USA and for solute types of diverse origin. While we observed C–Q patterns across geographically segregated catchment clusters, links between C–Q metrics (archetype and slope) and catchment and hydroclimatic attributes were frequently weak (Figure S8). Previous studies have also struggled to establish strong links (Ali et al., 2017; Musolff et al., 2015; Pohle et al., 2021; Wymore et al., 2017) at continental and annual scales due to several likely confounding factors. Seasonality and storm events affect C–Q relationships (Kincaid et al., 2020; Knapp et al., 2020; Knapp et al., 2022; Minaudo et al., 2019), and analyses of long-term low-frequency C–Q data can fail to capture seasonal and event-induced dynamics (Fazekas et al., 2020), especially when sample sizes are low. In addition to these temporal dynamics, study of C–Q relationships can fail to capture the influence of the spatial heterogeneity of solute source and discharge generating zones as streams reflect the integrated catchment response. Hot spots (McClain et al., 2003) or control points (Bernhardt et al., 2017) that disproportionately contribute streamflow and solute mass in heterogeneous catchments may not be well represented by catchment-average attributes. This study underscores the need for derived catchment-scale attributes that better characterize heterogeneity of solute sources—ideally developed from remote-sensing resources (e.g., land cover/land use and soil quality; Obade & Lal, 2013) to facilitate broad-scale ecohydrological studies. For example, Ebeling et al. (2021) computed two metrics to characterize the lateral component of source heterogeneity for diffuse NO₃ and phosphate sources in agricultural catchments using land cover maps and horizontal flow distances. To quantify vertical source heterogeneity, Ebeling et al. (2021) calculated the mean of the ratio between potential seepage NO₃ concentrations and estimated groundwater NO₃ concentrations.

Our inductive analysis can be used as a catalyst for ‘thought experiments’ to refine research questions or identify hypotheses for further testing using deductive investigative approaches at a regional or site scale. We provide example hypotheses generated from our study results.

4.3.1 | Hypothesis 1: Examine support for the shallow and deep hypothesis

The shallow and deep hypothesis states that the vertical profile of solute concentrations from shallow soils to deeper regolith and fractured bedrock shapes solute C–Q patterns in surface water at the catchment outlet (Zhi et al., 2019; Zhi & Li, 2020). Examples of prior support for this hypothesis include using a catchment-scale reactive transport model to link modelled subsurface water chemistry with stream chemistry (Wen et al., 2020; Zhi et al., 2019; Zhi & Li, 2020) and a direct comparison of solute concentration profiles in the subsurface to C–Q patterns (Stewart, Shanley, et al., 2022). Our study provides further support for the shallow and deep hypothesis at a continental scale, given the transition from dominantly dilution C–Q responses for geogenic solutes to dominantly enrichment C–Q responses for biogenic solutes (Figure 1).

Given our study’s support for the shallow and deep hypothesis (Section 4.2), we hypothesize that spatial heterogeneity (or distribution) of solutes in the vertical direction (from soils to bedrock) plays a more dominant role than source distribution in the lateral, landscape direction in determining C–Q patterns of surface water at the catchment outlet. Testing this hypothesis in catchments selected from representative geographic clusters using deductive approaches including process-based models could enhance understanding of these gradients. Studies in the vein of Zhi et al. (2019) and Stewart, Shanley, et al. (2022) could be replicated at additional observatories with sufficient vertical and lateral characterization across a range of geographic regions. If stream-water chemistry can be used to infer the vertical distribution and concentration levels of solutes as Stewart, Shanley, et al. (2022) demonstrated, this could be particularly useful because solute concentration data in soil and rocks are arduous and expensive to obtain.

4.3.2 | Hypothesis 2: Evaluate shifting threshold patterns in the face of disturbance

We hypothesize that threshold patterns will become more common for nutrients and bio-reactive solutes (e.g., carbon, nitrogen and phosphorus) with a changing climate. Given the observed nonstationarity in climate across the contiguous USA (Hirsch, 2011) and the regional differences in projected hydrometeorological trends (Hayhoe et al., 2007; IPCC, 2013), various regions may have been (and may continue to be) exposed to different magnitudes, frequencies and directionality of hydrologic drivers. Where winter snowfall turns to rain and snow retention declines, the role of snow in mediating water budgets, vegetation and fire (and their interactions) may change in profound ways (Gergel et al., 2017; Siirila-Woodburn et al., 2021).

Threshold patterns observed for the CAMELS-Chem C–Q record (1924–2020) may reflect catchment-specific transitions occurring within this period from dry to wet conditions (Knapp et al., 2022) or from dominantly biogeochemical processes to hydrologically dominated conditions (Musolff et al., 2015; Underwood et al., 2017). Additionally, aspects of catchment structure may be important in mediating the balance of hydrological and biogeochemical processes to govern stream-water chemistry dynamics at the catchment scale (Underwood et al., 2023) and contribute to system resilience in the face of disturbance. In a more deductive approach to investigation, data-driven methods (e.g., Bayesian hierarchical models) could be employed to test our hypothesis and identify shifting threshold positions by season, or by decade. Additionally, process-based models could facilitate a sensitivity analysis to estimate the water-budget effects from fire and changing snow storage and to quantify feedbacks between catchment structure and disturbance.

4.3.3 | Hypothesis 3: Examine role of hydrologic drivers in disproportionate threshold responses in Great Plains

We hypothesize that the combined physical structure and ecohydrological characteristics (Figure 3b) of Great Plains (cluster 3) catchments may lead to a greater frequency of threshold responses (C–Q archetypes C and F) for geogenic solutes as compared with other geographic regions. On an annualized basis, cluster 3 catchments have frequent low or no-flow states and/or groundwater-driven contributions to streamflow (Messenger et al., 2021), and the seasonal distribution of streamflow is different for this cluster than other clusters (Figure 3b) with higher discharge during the late spring and summer following snowmelt (Scott et al., 2019). Thus, the constant behaviour under low-flow conditions may reflect the extended contribution of geogenic-rich groundwater. Stream drying and intermittency may further contribute to high geogenic concentrations under low-flow conditions through evapo-concentration (Brooks & Lemon, 2007). Higher discharge in this region generally reflects inputs of snowmelt-derived water less enriched with geogenic solutes, resulting in the dilution behaviour observed under high-flow conditions (Scott et al., 2019; Warix et al., 2021). More deductive studies of these catchments via process-based models and higher frequency sampling may help determine the contribution of hydrologic drivers as controls on the observed threshold responses for geogenic solutes.

5 | CONCLUSIONS

We analysed C–Q relationships, including threshold relationships, for multiple solutes in minimally disturbed catchments across the contiguous USA. Resulting patterns provide key insights into controls on solute export. First, C–Q relationships with threshold responses were relatively rare compared with other studies in more intensively disturbed catchments, potentially indicating that urbanization and agricultural practices alter both solute pools and transport pathways and

ultimately C–Q relationships. However, more frequently paired C–Q observations that capture a larger range of flows at a given site may reveal a greater occurrence of threshold patterns resulting from variable seasonal or catchment wetness conditions (Burns et al., 2019; Knapp et al., 2022). Second, and similar to others, we found solute-specific ranges of C–Q metrics (Ebeling et al., 2021; Minaudo et al., 2019; Moatar et al., 2017; Zarnetske et al., 2018) across the contiguous USA, which support the shallow and deep hypothesis that C–Q relationships tend to be shaped by vertical solute–source distributions in catchments controlled by solute characteristics (e.g., solubility) that determine mobilization, transport, retention and ultimately solute export archetypes (Zhi et al., 2019; Zhi & Li, 2020). Third, both geographic location and catchment and hydroclimatic attributes contributed to C–Q relationship patterns for some solutes; however, we struggled to establish strong links to attributes, likely because the CAMELS attributes were developed at relatively coarse scales that do not sufficiently characterize solute–source variability in catchments. Fourth, our data-driven inductive approach revealed emergent patterns across solute groupings and regions of the contiguous USA. From process-guided interpretations of these patterns, we generated hypotheses that can be tested—along with others gleaned from patterns presented here—at discrete, representative catchments using deductive approaches to better understand the processes underlying solute export patterns. Finally, given these long-term C–Q relationships are from minimally disturbed catchments, our findings can be used as benchmarks for change in more disturbed catchments.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation's Critical Zone Network – Big Data Cluster (NSF EAR-2012123 and NSF EAR-2012080). D.W.K., D.M.R., K.L.U. and S.D.H. were also supported by NSF OIA-2033995. We would like to thank the Vermont EPSCoR program, especially Patrick Clemins, for the use and support of their high-performance computing resources. Additionally, Lauren Koenig, Heather Welch and three anonymous individuals provided reviews that greatly improved this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The version of the CAMELS-Chem dataset analysed here, accompanying R scripts, and supporting data files are available via HydroShare (Kincaid & Underwood, 2024; <https://doi.org/10.4211/hs.eddb06e91a914618a89a63bb2c2774e0>).

REFERENCES

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21(10), 5293–5313. <https://doi.org/10.5194/hess-21-5293-2017>

- Alfaro, M. A., Jarvis, S. C., & Gregory, P. J. (2004). Factors affecting potassium leaching in different soils. *Soil Use and Management*, 20(2), 182–189. <https://doi.org/10.1111/j.1475-2743.2004.tb00355.x>
- Ali, G., Wilson, H., Elliott, J., Penner, A., Haque, A., Ross, C., & Rabie, M. (2017). Phosphorus export dynamics and hydrobiogeochemical controls across gradients of scale, topography and human impact. *Hydrological Processes*, 31(18), 3130–3145. <https://doi.org/10.1002/hyp.11258>
- Basu, N. B., Destouni, G., Jawitz, J. W., Thompson, S. E., Loukinova, N. V., Darracq, A., Zanardo, S., Yaeger, M., Sivapalan, M., Rinaldo, A., & Rao, P. S. C. (2010). Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophysical Research Letters*, 37(23), L23404. <https://doi.org/10.1029/2010gl045168>
- Bende-Michl, U., Verburg, K., & Cresswell, H. P. (2013). High-frequency nutrient monitoring to infer seasonal patterns in catchment source availability, mobilisation and delivery. *Environmental Monitoring and Assessment*, 185(11), 9191–9219. <https://doi.org/10.1007/s10661-013-3246-8>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berner, E. K., & Berner, R. A. (2012). *Global environment: Water, air, and geochemical cycles*. Princeton University Press.
- Bernhardt, E. S., Blaszczak, J. R., Ficken, C. D., Fork, M. L., Kaiser, K. E., & Seybold, E. C. (2017). Control points in ecosystems: Moving beyond the hot spot hot moment concept. *Ecosystems*, 20(4), 665–682. <https://doi.org/10.1007/s10021-016-0103-y>
- Beven, K. J. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, 4(2), 203–213. <https://doi.org/10.5194/hess-4-203-2000>
- Botter, M., Li, L., Hartmann, J., Burlando, P., & Faticchi, S. (2020). Depth of solute generation is a dominant control on concentration-discharge relations. *Water Resources Research*, 56(8), e2019WR026695. <https://doi.org/10.1029/2019WR026695>
- Boyer, E. W., Hornberger, G. M., Bencala, K. E., & McKnight, D. M. (1997). Response characteristics of DOC flushing in an alpine catchment. *Hydrological Processes*, 11(12), 1635–1647. [https://doi.org/10.1002/\(sici\)1099-1085\(19971015\)11:12<1635::aid-hyp494>3.0.co;2-h](https://doi.org/10.1002/(sici)1099-1085(19971015)11:12<1635::aid-hyp494>3.0.co;2-h)
- Brooks, P. D., & Lemon, M. M. (2007). Spatial variability in dissolved organic matter and inorganic nitrogen concentrations in a semiarid stream, San Pedro River, Arizona. *Journal of Geophysical Research: Biogeosciences*, 112(G3), G03S05. <https://doi.org/10.1029/2006jg000262>
- Burns, D. A., Pellerin, B. A., Miller, M. P., Capel, P. D., Tesoriero, A. J., & Duncan, J. M. (2019). Monitoring the riverine pulse: Applying high-frequency nitrate data to advance integrative understanding of biogeochemical and hydrological processes. *Wiley Interdisciplinary Reviews: Water*, 6(4), e1348. <https://doi.org/10.1002/wat2.1348>
- Cain, M. R., Woo, D. K., Kumar, P., Keefer, L., & Ward, A. S. (2022). Antecedent conditions control thresholds of tile-runoff generation and nitrogen export in intensively managed landscapes. *Water Resources Research*, 58(2), e2021WR030507. <https://doi.org/10.1029/2021wr030507>
- Chorover, J., Derry, L. A., & McDowell, W. H. (2017). Concentration-discharge relations in the critical zone: Implications for resolving critical zone structure, function, and evolution. *Water Resources Research*, 53(11), 8654–8659. <https://doi.org/10.1002/2017wr021111>
- Creed, I. F., McKnight, D. M., Pellerin, B. A., Green, M. B., Bergamaschi, B. A., Aiken, G. R., Burns, D. A., Findlay, S. E. G., Shanley, J. B., Striegl, R. G., Aulenbach, B. T., Clow, D. W., Laudon, H., McGlynn, B. L., McGuire, K. J., Smith, R. A., & Stackpoole, S. M. (2015). The river as a chemostat: Fresh perspectives on dissolved organic matter flowing down the river continuum. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(8), 1272–1285. <https://doi.org/10.1139/cjfas-2014-0400>

- Diamond, J. S., & Cohen, M. J. (2018). Complex patterns of catchment solute–discharge relationships for coastal plain rivers. *Hydrological Processes*, 32(3), 388–401. <https://doi.org/10.1002/hyp.11424>
- Dupas, R., Musolff, A., Jawitz, J. W., Rao, P. S. C., Jäger, C. G., Fleckenstein, J. H., Rode, M., & Borchardt, D. (2017). Carbon and nutrient export regimes from headwater catchments to downstream reaches. *Biogeosciences*, 14(18), 4391–4407. <https://doi.org/10.5194/bg-14-4391-2017>
- Ebeling, P., Kumar, R., Weber, M., Knoll, L., Fleckenstein, J. H., & Musolff, A. (2021). Archetypes and controls of riverine nutrient export across German catchments. *Water Resources Research*, 57(4), e2020WR028134. <https://doi.org/10.1029/2020wr028134>
- Fazekas, H. M., Wymore, A. S., & McDowell, W. H. (2020). Dissolved organic carbon and nitrate concentration–discharge behavior across scales: Land use, excursions, and misclassification. *Water Resources Research*, 56(8), e2019WR027028. <https://doi.org/10.1029/2019wr027028>
- Gall, H. E., Park, J., Harman, C. J., Jawitz, J. W., & Rao, P. S. C. (2013). Landscape filtering of hydrologic and biogeochemical responses in managed catchments. *Landscape Ecology*, 28(4), 651–664. <https://doi.org/10.1007/s10980-012-9829-x>
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Chapman & Hall/CRC.
- Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. *Climatic Change*, 141(2), 287–299. <https://doi.org/10.1007/s10584-017-1899-y>
- Godsey, S. E., Hartmann, J., & Kirchner, J. W. (2019). Catchment chemostasis revisited: Water quality responds differently to variations in weather and climate. *Hydrological Processes*, 33(24), 3056–3069. <https://doi.org/10.1002/hyp.13554>
- Godsey, S. E., Kirchner, J. W., & Clow, D. W. (2009). Concentration–discharge relationships reflect chemostatic characteristics of US catchments. *Hydrological Processes*, 23(13), 1844–1864. <https://doi.org/10.1002/hyp.7315>
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27(4), 857. <https://doi.org/10.2307/2528823>
- Hayhoe, K., Wake, C. P., Huntington, T. G., Luo, L., Schwartz, M. D., Sheffield, J., Wood, E., Anderson, B., Bradbury, J., DeGaetano, A., Troy, T. J., & Wolfe, D. (2007). Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dynamics*, 28(4), 381–407. <https://doi.org/10.1007/s00382-006-0187-8>
- Herndon, E. M., Dere, A. L., Sullivan, P. L., Norris, D., Reynolds, B., & Brantley, S. L. (2015). Landscape heterogeneity drives contrasting concentration–discharge relationships in shale headwater catchments. *Hydrology and Earth System Sciences*, 19(8), 3333–3347. <https://doi.org/10.5194/hess-19-3333-2015>
- Hirsch, R. M. (2011). A perspective on nonstationarity and water management. *Journal of the American Water Resources Association*, 47(3), 436–446. <https://doi.org/10.1111/j.1752-1688.2011.00539.x>
- Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, 45(7), 1–47. <https://doi.org/10.18637/jss.v045.i07>
- Hornberger, G. M., Bencala, K. E., & McKnight, D. M. (1994). Hydrological controls on dissolved organic carbon during snowmelt in the Snake River near Montezuma, Colorado. *Biogeochemistry*, 25(3), 147–165. <https://doi.org/10.1007/bf00024390>
- IPCC. (2013). *Climate change 2013: The physical science basis*. Cambridge University Press.
- Johnson, N. M., Likens, G. E., Bormann, F. H., Fisher, D. W., & Pierce, R. S. (1969). A working model for the variation in stream water chemistry at the Hubbard Brook Experimental Forest, New Hampshire. *Water Resources Research*, 5(6), 1353–1363. <https://doi.org/10.1029/wr005i006p01353>
- Kay, M. (2021). tidybayes: Tidy data and geoms for Bayesian models. R package version 3.0.1. <https://doi.org/10.5281/zenodo.1308151>
- Kincaid, D., & Underwood, K. (2024). *Code and data for CAMELS-Chem concentration-discharge analysis* [Data set]. HydroShare. <https://doi.org/10.4211/hs.eddb06e91a914618a89a63bb2c2774e0>
- Kincaid, D. W., Seybold, E. C., Adair, E. C., Bowden, W. B., Perdrial, J. N., Vaughan, M. C. H., & Schroth, A. W. (2020). Land use and season influence event-scale nitrate and soluble reactive phosphorus exports and export stoichiometry from headwater catchments. *Water Resources Research*, 56(10), e2020WR027361. <https://doi.org/10.1029/2020wr027361>
- Knapp, J. L. A., Von, F. J., Studer, B., Kiewiet, L., & Kirchner, J. W. (2020). Concentration–discharge relationships vary among hydrological events, reflecting differences in event characteristics. *Hydrology and Earth System Sciences*, 24(5), 2561–2576. <https://doi.org/10.5194/hess-24-2561-2020>
- Knapp, J. L. A., Li, L., & Musolff, A. (2022). Hydrologic connectivity and source heterogeneity control concentration–discharge relationships. *Hydrological Processes*, 36(9), e14683. <https://doi.org/10.1002/hyp.14683>
- Levin, S. A. (1992). The problem of pattern and scale in ecology: The Robert H MacArthur Award Lecture. *Ecology*, 73(6), 1943–1967. <https://doi.org/10.2307/1941447>
- Li, L., Stewart, B., Zhi, W., Sadayappan, K., Ramesh, S., Kerins, D., Sterle, G., Harpold, A., & Perdrial, J. (2022). The climate control on river chemistry. *Earth's Future*, 10, e2021EF002603. <https://doi.org/10.1029/2021ef002603>
- Likens, G. E., Driscoll, C. T., Buso, D. C., Siccama, T. G., Johnson, C. E., Lovett, G. M., Ryan, D. F., Fahey, T., & Reiners, W. A. (1994). The biogeochemistry of potassium at Hubbard Brook. *Biogeochemistry*, 25(2), 61–125. <https://doi.org/10.1007/bf00000881>
- Lindeløv, J. K. (2020). mcp: An R package for regression with multiple change points. <https://doi.org/10.31219/osf.io/fzqvx>
- Lins, H. F. (2012). USGS hydro-climatic data network 2009 (HCDN-2009). U.S. Geological Survey Fact Sheet 2012–3047. Available from <https://pubs.usgs.gov/fs/2012/3047/>
- Ma, Z., Guan, K., Peng, B., Sivapalan, M., Li, L., Pan, M., Zhou, W., Warner, R., & Zhang, J. (2022). Agricultural nitrate export patterns shaped by crop rotation and tile drainage. *Water Research*, 229, 119468. <https://doi.org/10.1016/j.watres.2022.119468>
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2021). cluster: Cluster analysis basics and extensions. R package version 2.1.2. Available from <https://cran.r-project.org/package=cluster>
- Marinos, R. E., Meter, K. J. V., & Basu, N. B. (2020). Is the river a chemostat?: Scale versus land use controls on nitrate concentration–discharge dynamics in the Upper Mississippi River Basin. *Geophysical Research Letters*, 47(16), e2020GL087051. <https://doi.org/10.1029/2020gl087051>
- Mayer, B., Shanley, J. B., Bailey, S. W., & Mitchell, M. J. (2010). Identifying sources of stream water sulfate after a summer drought in the Sleepers River watershed (Vermont, USA) using hydrological, chemical, and isotopic techniques. *Applied Geochemistry*, 25(5), 747–754. <https://doi.org/10.1016/j.apgeochem.2010.02.007>
- McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., Hart, S. C., Harvey, J. W., Johnston, C. A., Mayorga, E., Pinay, G., & Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, 6(4), 301–312. <https://doi.org/10.1007/s10021-003-0161-9>
- Messenger, M. L., Lehner, B., Cockburn, C., Lamouroux, N., Pella, H., Snelder, T., Tockner, K., Trautmann, T., Watt, C., & Datry, T. (2021). Global prevalence of non-perennial rivers and streams. *Nature*, 594(7863), 391–397. <https://doi.org/10.1038/s41586-021-03565-5>
- Meybeck, M., & Moatar, F. (2012). Daily variability of river concentrations and fluxes: Indicators based on the segmentation of the rating curve. *Hydrological Processes*, 26(8), 1188–1207. <https://doi.org/10.1002/hyp.8211>

- Meyer, D., Zeileis, A., & Hornik, K. (2021). vcd: Visualizing categorical data. R package version 1.4-9.
- Minaudo, C., Dupas, R., Gascuel-Oudoux, C., Roubeix, V., Danis, P.-A., & Moatar, F. (2019). Seasonal and event-based concentration-discharge relationships to identify catchment controls on nutrient export regimes. *Advances in Water Resources*, 131(8), 103379. <https://doi.org/10.1016/j.advwatres.2019.103379>
- Moatar, F., Abbott, B. W., Minaudo, C., Curie, F., & Pinay, G. (2017). Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. *Water Resources Research*, 53(2), 1270–1287. <https://doi.org/10.1002/2016wr019635>
- Musolff, A., Fleckenstein, J. H., Rao, P. S. C., & Jawitz, J. W. (2017). Emergent archetype patterns of coupled hydrologic and biogeochemical responses in catchments. *Geophysical Research Letters*, 44(9), 4143–4151. <https://doi.org/10.1002/2017gl072630>
- Musolff, A., Schmidt, C., Selle, B., & Fleckenstein, J. H. (2015). Catchment controls on solute export. *Advances in Water Resources*, 86, 133–146. <https://doi.org/10.1016/j.advwatres.2015.09.026>
- National Science Foundation (NSF). (2018). Bridging the atom-to-global scale gap. Available from <https://nsf2026imgallery.skild.com/entries/bridging-the-atom-to-global-scale-gap>
- Neal, C., Robson, A., & Smith, C. J. (1990). Acid neutralization capacity variations for the Hafren forest stream, mid-Wales: Inferences for hydrological processes. *Journal of Hydrology*, 121(1–4), 85–101. [https://doi.org/10.1016/0022-1694\(90\)90226-n](https://doi.org/10.1016/0022-1694(90)90226-n)
- Obade, V. D. P., & Lal, R. (2013). Assessing land cover and soil quality by remote sensing and geographical information systems (GIS). *CATENA*, 104, 77–92. <https://doi.org/10.1016/j.catena.2012.10.014>
- Omerik, J. M. (1987). Ecoregions of the conterminous United States. *Annals of the Association of American Geographers*, 77(1), 118–125. <https://doi.org/10.1111/j.1467-8306.1987.tb00149.x>
- Pohle, I., Baggaley, N., Palarea-Albaladejo, J., Stutter, M., & Glendell, M. (2021). A framework for assessing concentration-discharge catchment behavior from low-frequency water quality data. *Water Resources Research*, 57(9), e2021WR029692. <https://doi.org/10.1029/2021wr029692>
- R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Raymond, P. A., Saiers, J. E., & Sobczak, W. V. (2016). Hydrological and biogeochemical controls on watershed dissolved organic matter transport: Pulse-shunt concept. *Ecology*, 97(1), 5–16. <https://doi.org/10.1890/14-1684.1>
- Rose, L. A., Karwan, D. L., & Godsey, S. E. (2018). Concentration–discharge relationships describe solute and sediment mobilization, reaction, and transport at event and longer timescales. *Hydrological Processes*, 32(18), 2829–2844. <https://doi.org/10.1002/hyp.13235>
- Sardans, J., & Peñuelas, J. (2015). Potassium stoichiometry and global change. *Global Ecology and Biogeography*, 24(3), 261–275. <https://doi.org/10.1111/geb.12259>
- Scott, D. T., Gomez-Velez, J. D., Jones, C. N., & Harvey, J. W. (2019). Floodplain inundation spectrum across the United States. *Nature Communications*, 10(1), 5194. <https://doi.org/10.1038/s41467-019-13184-4>
- Seibert, J., Grabs, T., Köhler, S., Laudon, H., Winterdahl, M., & Bishop, K. (2009). Linking soil- and stream-water chemistry based on a Riparian Flow-Concentration Integration Model. *Hydrology and Earth System Sciences*, 13(12), 2287–2297. <https://doi.org/10.5194/hess-13-2287-2009>
- Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., Collins, W. D., & Kaatz, L. (2021). A low-to-no snow future and its impacts on water resources in the western United States. *Nature Reviews Earth & Environment*, 2(11), 800–819. <https://doi.org/10.1038/s43017-021-00219-y>
- Speir, S. L., Rose, L. A., Blaszczyk, J. R., Kincaid, D. W., Fazekas, H. M., Webster, A. J., Wolford, M. A., Shogren, A. J., & Wymore, A. S. (2024). Catchment concentration–discharge relationships across temporal scales: A review. *Wiley Interdisciplinary Reviews: Water*, 11(2), e1702. <https://doi.org/10.1002/wat2.1702>
- Sterle, G., Perdrial, J., Kincaid, D. W., Underwood, K. L., Rizzo, D. M., Haq, I. U., Li, L., Lee, B. S., Adler, T., Wen, H., Middleton, H., & Harpold, A. A. (2024). CAMELS-Chem: Augmenting CAMELS (Catchment Attributes and Meteorology for Large-sample Studies) with atmospheric and stream water chemistry data. *Hydrology and Earth System Sciences*, 28(3), 611–630. <https://doi.org/10.5194/hess-28-611-2024>
- Stewart, B., Shanley, J. B., Kirchner, J. W., Norris, D., Adler, T., Bristol, C., Harpold, A. A., Perdrial, J. N., Rizzo, D. M., Sterle, G., Underwood, K. L., Wen, H., & Li, L. (2022). Streams as mirrors: Reading subsurface water chemistry from stream chemistry. *Water Resources Research*, 58(1), e2021WR029931. <https://doi.org/10.1029/2021wr029931>
- Stewart, B., Zhi, W., Sadayappan, K., Sterle, G., Harpold, A., & Li, L. (2022). Soil CO₂ controls short-term variation but climate regulates long-term mean of riverine inorganic carbon. *Global Biogeochemical Cycles*, 36(8), e2022GB007351. <https://doi.org/10.1029/2022gb007351>
- Thompson, S. E., Basu, N. B., Lascrain, J., Aubeneau, A., & Rao, P. S. C. (2011). Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients. *Water Resources Research*, 47(10), W00J05. <https://doi.org/10.1029/2010wr009605>
- Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, 18(4), 267–276. <https://doi.org/10.1007/bf02289263>
- Underwood, K. L., Rizzo, D. M., Hanley, J. P., Sterle, G., Harpold, A., Adler, T., Li, L., Wen, H., & Perdrial, J. N. (2023). Machine-learning reveals equifinality in drivers of stream DOC concentration at continental scales. *Water Resources Research*, 59(3), e2021WR030551. <https://doi.org/10.1029/2021wr030551>
- Underwood, K. L., Rizzo, D. M., Schroth, A. W., & Dewoolkar, M. M. (2017). Evaluating spatial variability in sediment and phosphorus concentration-discharge relationships using Bayesian inference and self-organizing maps. *Water Resources Research*, 53(12), 10293–10316. <https://doi.org/10.1002/2017wr021353>
- U.S. Geological Survey. (2023). USGS water data for the nation: U.S. Geological Survey National Water Information System database. <https://doi.org/10.5066/F7P55KJN>
- Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P., Paananen, T., & Gelman, A. (2020). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version. Available from <https://mc-stan.org/loo/>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vidon, P. G. F., & Hill, A. R. (2004). Landscape controls on nitrate removal in stream riparian zones. *Water Resources Research*, 40(3), 707–714. <https://doi.org/10.1029/2003wr002473>
- Ward, A. S., Packman, A., Bernal, S., Brekenfeld, N., Drummond, J., Graham, E., Hannah, D. M., Klauf, M., Krause, S., Kurz, M., Li, A., Lupon, A., Mao, F., Roca, M. E. M., Ouellet, V., Royer, T. V., Stegen, J. C., & Zarnetske, J. P. (2022). Advancing river corridor science beyond disciplinary boundaries with an inductive approach to catalyse hypothesis generation. *Hydrological Processes*, 36(4), e14540. <https://doi.org/10.1002/hyp.14540>
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Warix, S. R., Godsey, S. E., Lohse, K. A., & Hale, R. L. (2021). Influence of groundwater and topography on stream drying in semi-arid headwater streams. *Hydrological Processes*, 35(5), e14185. <https://doi.org/10.1002/hyp.14185>

- Wen, H., Perdrial, J., Abbott, B. W., Bernal, S., Dupas, R., Godsey, S. E., Harpold, A., Rizzo, D., Underwood, K., Adler, T., Sterle, G., & Li, L. (2020). Temperature controls production but hydrology regulates export of dissolved organic carbon at the catchment scale. *Hydrology and Earth System Sciences*, 24(2), 945–966. <https://doi.org/10.5194/hess-24-945-2020>
- Wymore, A. S., Brereton, R. L., Ibarra, D. E., Maher, K., & McDowell, W. H. (2017). Critical zone structure controls concentration-discharge relationships and solute generation in forested tropical montane watersheds. *Water Resources Research*, 53(7), 6279–6295. <https://doi.org/10.1002/2016wr020016>
- Wymore, A. S., Fazekas, H. M., & McDowell, W. H. (2021). Quantifying the frequency of synchronous carbon and nitrogen export to the river network. *Biogeochemistry*, 152(1), 1–12. <https://doi.org/10.1007/s10533-020-00741-z>
- Wymore, A. S., Larsen, W., Kincaid, D. W., Underwood, K. L., Fazekas, H. M., McDowell, W. H., Murray, D. S., Shogren, A. J., Speir, S. L., & Webster, A. J. (2023). Revisiting the origins of the power-law analysis for the assessment of concentration-discharge relationships. *Water Resources Research*, 59(8), e2023WR034910. <https://doi.org/10.1029/2023wr034910>
- Zarnetske, J. P., Bouda, M., Abbott, B. W., Saiers, J., & Raymond, P. A. (2018). Generality of hydrologic transport limitation of watershed organic carbon flux across ecoregions of the United States. *Geophysical Research Letters*, 45(21), 11702–11711. <https://doi.org/10.1029/2018gl080005>
- Zeileis, A., Meyer, D., & Hornik, K. (2007). Residual-based shadings for visualizing (conditional) independence. *Journal of Computational and Graphical Statistics*, 16(3), 507–525. <https://doi.org/10.1198/106186007x237856>
- Zhi, W., & Li, L. (2020). The shallow and deep hypothesis: Subsurface vertical chemical contrasts shape nitrate export patterns from different land uses. *Environmental Science & Technology*, 54(19), 11915–11928. <https://doi.org/10.1021/acs.est.0c01340>
- Zhi, W., Li, L., Dong, W., Brown, W., Kaye, J., Steefel, C., & Williams, K. H. (2019). Distinct source water chemistry shapes contrasting concentration-discharge patterns. *Water Resources Research*, 55(5), 4233–4251. <https://doi.org/10.1029/2018wr024257>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kincaid, D. W., Underwood, K. L., Hamshaw, S. D., Li, L., Seybold, E. C., Stewart, B., Rizzo, D. M., Haq, I. U., & Perdrial, J. N. (2024). Solute export patterns across the contiguous USA. *Hydrological Processes*, 38(6), e15197. <https://doi.org/10.1002/hyp.15197>