Network Analysis Reveals Multi-Scale Controls on Streamwater Chemistry


1Virginia Water Resources Research Center and Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA 24061;
2U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Cascadia Field Station and School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195;
3Cary Institute of Ecosystem Studies, Millbrook, NY 12545 and Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269;
4Cary Institute of Ecosystem Studies, Hubbard Brook Forest Station, North Woodstock, NH 03262;
5Division of Biological Sciences, The University of Montana, Missoula, MT 59812;
6U.S. Forest Service, Northern Research Station, Hubbard Brook Experimental Forest, North Woodstock, NH 03262

*Corresponding authors


Author contributions: C.E.T., K.J.M., and G.E.L. designed the research; D.C.B. conducted fieldwork and laboratory analyses; C.E.T. and K.J.M. analyzed data and C.E.T. performed the statistical analysis; and K.J.M., C.E.T., G.E.L., W.H.L., and S.W.B. wrote the paper.

The authors declare no conflict of interest.

This draft manuscript is distributed solely for purposes of scientific peer review. Its content is deliberative and pre-decisional, so it must not be disclosed or released by reviewers. Because the manuscript has not yet been approved for publication by the U.S. Geological Survey (USGS), it does not represent any official USGS finding or policy.
Abstract

By coupling synoptic data from a basin-wide assessment of streamwater chemistry with network-based geostatistical analysis, we show that spatial processes differentially affect biogeochemical condition and pattern across a headwater stream network. We analyzed a high-resolution dataset consisting of 664 water samples collected every 100 m throughout 32 tributaries in an entire fifth-order stream network. These samples were analyzed for an exhaustive suite of chemical constituents. The fine grain and broad extent of this study design allowed us to quantify spatial patterns over a range of scales using empirical semivariograms that explicitly incorporated network topology. Here, we show that spatial structure, as determined by the characteristic shape of the semivariograms, differed both among chemical constituents and by spatial relationship (flow-connected, flow-unconnected, or Euclidean). Spatial structure was apparent at either a single scale or at multiple nested scales, suggesting separate processes operating simultaneously within the stream network and surrounding terrestrial landscape. Expected patterns of spatial dependence for flow-connected relationships (e.g., increasing homogeneity with downstream distance) occurred for some chemical constituents (e.g., dissolved organic carbon, sulfate, and aluminum) but not for others (e.g., nitrate, sodium). By comparing semivariograms for the different chemical constituents and spatial relationships, we were able to separate effects on streamwater chemistry of (1) fine- versus broad-scale processes and (2) in-stream processes versus landscape controls. These findings provide novel insight on the hierarchical scaling of local, longitudinal, and landscape processes that drive biogeochemical patterns in stream networks.

Key words: biogeochemistry, spatial pattern, stream network, connectivity, scale

Significance Statement

Headwater streams are important sources of water for downstream ecosystems and human communities. These streams comprise the vast majority of stream and river kilometers in watersheds and affect regional water quality. However, the actual spatial variation of water quality in headwater streams is often unknown. Our study uses an unusually high-resolution spatial dataset from a headwater stream network and employs a new statistical tool to objectively describe spatial patterns of streamwater chemistry within a stream network. This approach provides new insights on how flowing water interacts with vegetation, soil, and geologic materials in the surrounding landscape. Application of this method may help to identify factors impairing water quality and to inform strategies for protecting aquatic ecosystems.
Introduction
Spatial heterogeneity of ecosystems has been a focus of landscape ecology for over two decades, but the linkages between these patterns and underlying processes are still poorly understood (1-3). Quantifying these pattern-process links is largely a problem of scale. Specifically, it is difficult to perform experiments at the landscape scale and measure responses over the range of spatial and temporal scales commensurate with the processes of interest (4, 5).

This problem of scale limits our understanding of both terrestrial and freshwater ecosystems. Effects of landscape pattern on ecosystem response can be evaluated at stream outlets using biogeochemical signals that integrate physical and biological conditions of the catchment (6, 7). However, the spatial complexity of biogeochemical patterns and processes within stream networks has not been fully investigated because it is difficult to quantify such patterns at a grain and extent sufficient for examining spatial heterogeneity and processes across scales (8). Quantifying this variability and linking fine- and broad-scale patterns and processes within the branched topology of stream networks is essential for understanding aquatic ecosystem function and aquatic-terrestrial ecosystem connections, but requires new conceptual and methodological approaches (9, 10).

Major advances in understanding biogeochemical fluxes and cycles in rivers and streams have resulted from increased recognition of how spatial heterogeneity and network topology reflect land–water interactions (e.g., 11, 12). However, our understanding of biogeochemical processes in stream networks is still limited to small-scale experiments (e.g., 13), often with limited spatial extent or replication, and large-scale correlative models (14). Fine-grained observations at intermediate scales (e.g., 1 to 10 km²) may be especially powerful for advancing understanding of complex aquatic and terrestrial effects on biogeochemical fluxes throughout stream networks (15-17).

Studies quantifying streamwater chemistry in a spatially intensive manner at intermediate scales have revealed a high degree of spatial structure that cannot be explained by current models of biogeochemical processes (11, 18). Specifically, these results show that traditional, continuum-based models—where conditions are regulated primarily by upstream processes, and thus exhibit gradual downstream gradients—are insufficient for describing the true spatial complexity of biogeochemical patterns and processes in stream networks. This unfamiliar ground between fine and coarse scales of understanding is the crux of field-based science, in which the “preferred modes of explanation…appear to be systematically related to customary human scales of perception of the world” (19). Likewise, obtaining a bird’s-eye view of biogeochemical patterns at fine to coarse scales may be crucial for advancing ecosystem science and explaining the spatial complexity of streamwater chemistry within landscapes.

Recent developments in geostatistical modeling provide a valuable new perspective on stream networks by revealing hydrological and ecological patterns in a spatially continuous manner (20, 21). To date, the relatively few sample points required to generate spatial interpolations have fueled the popularity of these models. However, the increasing use of network-based geostatistical techniques underscores a need to understand the processes from which these
patterns arise, or, more broadly, to elucidate ecosystem processes from spatial patterns and develop new hypotheses about system function (22). Recent theoretical and empirical approaches show that this is possible using empirical semivariograms and synoptic sampling (e.g., 23). Specifically, the combination of spatial analysis and synoptic sampling allows one to visualize how patterns occur across different scales, while providing the empirical foundation needed to identify the processes that give rise to those patterns. Geostatistics have only recently been used to describe spatial patterns throughout stream networks (24, 25), even though these tools have long been used to quantify spatial structure in terrestrial ecology (26).

We apply geostatistical techniques to an unusually high-resolution synoptic dataset of streamwater chemistry collected throughout the Hubbard Brook Valley in New Hampshire to explore the spatial structure of biogeochemical patterns at multiple scales (18). The dataset consisted of 664 water samples collected over a three-month period every 100 m throughout all 32 tributaries of the 3600 ha, fifth-order stream network of the Hubbard Brook Valley. We show previously undescribed patterns of spatial dependence based on three spatial relationships, revealing biogeochemical determinants occurring across scales, both within the stream network and surrounding catchment. Stream network patterns were defined by two spatial relationships: flow-connected and flow-unconnected (sensu 20, 21). The straight-line distance between two points defines Euclidean relationships. Flow-connected and unconnected network relationships describe distances along the stream network and were considered “connected” if water flows from one site to another. Thus, all points downstream of other points on the stream network were considered connected, but points upstream of tributary junctions that do not share flow were considered “unconnected.”

Empirical semivariograms based on these three spatial relationships suggest the importance of different drivers of spatial variability in streamwater chemistry at multiple scales, e.g., fine (<1500 m) and broad scales (>3000 m) (Figure 1). For example, semivariograms of flow-connected relationships indicate whether downstream flow and longitudinal transport exert a dominant control on streamwater chemistry by showing the level of autocorrelation between flow-connected samples. Likewise, semivariograms of flow-unconnected relationships provide information about the similarity/dissimilarity of tributary branches due to influences of landscape properties (e.g., soils or geology). Semivariograms of streamwater chemistry using Euclidean relationships reveal interactions or lateral connectivity between the stream network and the landscape. Therefore, both Euclidean and flow-unconnected network relationships provide information on how the landscape influences patterns of streamwater chemistry within a single catchment/network, whereas a flow-connected relationship largely describes the effect of hydrologic transport and upstream spatial dependence.

Extensive work in the Hubbard Brook Ecosystem Study (HBES) over the last five decades provides the temporal context for understanding biogeochemical processes and landscape change through ecosystem change revealed by long-term research (27, 28). The current study aims to provide a spatial context (29) for interpreting how biogeochemical patterns observed from sparse fixed sites (e.g., outlets of experimental watersheds) fit within the larger stream network. We expect spatial dependence of streamwater chemistry to be structured by flow directionality and network topology, especially for constituents that are not strongly biologically cycled in headwater streams (e.g., base cations, Cl\(^-\), SO\(_4^{2-}\)). However, patchiness longitudinally in the
stream network and across the landscape (i.e., by Euclidean distances) may arise due to the local
effects of landscape features such as seeps and springs, and variation in vegetation, soil, and
geologic materials. Our objectives were to (1) quantify spatial heterogeneity in streamwater
chemistry at multiple scales within the stream network, (2) compare patterns of streamwater
chemistry using different spatial relationships within the stream network and across the
landscape (i.e., using network and Euclidean relationships), and (3) evaluate this approach for
linking biogeochemical patterns and processes by identifying potential drivers of spatial patterns
in streamwater chemistry that bridge scales from tributaries, to the main stem, and throughout the
entire Hubbard Brook Valley.

Results
Spatial structuring in streamwater chemistry
Streamwater chemistry patterns throughout the Hubbard Brook Valley exhibited spatial
patchiness and gradients at multiple scales (see supplemental information [Figure S1] for a
complete set of distribution maps for all chemical constituents). However, it is difficult to
systematically compare these patterns visually due to the spatial complexity of the data. For
example, patterns of dissolved organic carbon (DOC), sodium, dissolved silica, and specific
conductance were visibly different (Figure 2), but the extent and typology of these differences
cannot be quantified based on visual inspection alone. In contrast, analysis with empirical
semivariograms revealed a high degree of spatial structuring in streamwater chemistry, as
indicated by the overall shape of the semivariograms (Figures 3 and S2). Specifically, some of
the semivariograms showed spatial structure at predominantly one scale of variation (<1500 m)
(cf. Figure 1c, flow-connected semivariograms in Figure 3), whereas other semivariograms
varied substantially among spatial relationships and exhibited spatial structure at multiple scales
(cf. Figure 1d, Figure 3). Semivariograms of the 16 chemical constituents exhibited several of
the theoretical spatial structures described in Figure 1 (Table 1, Figures 3 and S2): (1)
nonstructured, or poorly defined structure (NH$_4^+$, NO$_3^-$, PO$_4^{3-}$; cf. Figure 1a); (2) single-scale
structure (many examples; cf. Figure 1c); and (3) nested structure (many examples; cf. Figure
1d). Scales of variation in streamwater chemistry ranged from 500 to >6000 m for the variables
that showed spatial structure (Table 1). The finest and broadest scales of variation were
associated with semivariograms of the nested type, for which Euclidean and flow-unconnected
relationships had the finest scales of variation, and flow-unconnected relationships had the
broadest scales of variation. Scales of variation for flow-connected network relationships were
less variable compared to the other spatial relationships.

Effects of network connectivity on patterns of spatial dependence
Patterns of spatial dependence in streamwater chemistry varied not only among chemical
constituents but also depending on how connectivity, or relationships among sample points, was
defined in space. For example, semivariograms of specific conductance differed in shape and
overall variance (i.e., height of the curve with respect to the y-axis) based on the way the spatial
relationship was calculated between sample points (i.e., flow-connected, flow-unconnected, and
Euclidean spatial relationships) (Figure 3d). Specifically, the shape of the flow-connected
semivariogram of specific conductance indicated the presence of spatial structure at one scale
(Figure 3d), whereas the semivariogram of flow-unconnected relationships exhibited pronounced
patterns associated with nested spatial structure at two scales. The semivariogram based on the
Euclidean spatial relationship also exhibited a nested pattern, but was not as clearly defined as
the flow-unconnected semivariogram. Overall variance was generally highest for the semivariogram based on Euclidean relationships and lowest for the flow-connected semivariograms.

The differences described above for semivariograms of specific conductance generally applied to the majority of chemical constituents that exhibited spatial structure (i.e., monomeric aluminum [Al₃], Cl⁻, dissolved inorganic carbon [DIC], DOC, K⁺, dissolved Si, and SO₄²⁻) (Figure S2), but there were some notable exceptions. For example, semivariograms of acid-neutralizing capacity (ANC), Ca²⁺, and K⁺ were also similar, but only for flow-connected and Euclidean spatial relationships with distances less than about 4500 m (Figure S2). Semivariograms of Mg²⁺ and Na⁺ exhibited an analogous pattern but at shorter distances (~1500 m). The nugget value (i.e., discontinuity at the origin) (e.g., defined in Figure 1c) of semivariograms based on Euclidean and flow-connected relationships were close to zero for all chemical constituents except for DIC (Figures 3 and S2), whereas nugget-variance for flow-unconnected semivariograms were substantially greater than zero for all chemical constituents except Al₃ and K⁺.

Discussion

Understanding controls on spatial patterns of streamwater chemistry and developing tools for predicting spatial variation of stream chemistry in headwaters is important for managing water quality in downstream systems. Headwater streams comprise the vast majority of stream length in watersheds (30) and perform critical functions for downstream ecosystems, but are still considered *aqua incognita* in hydrology and ecology (31, 32). We found complex variability in spatial patterns of streamwater chemistry (multi-scale structure) across the Hubbard Brook Valley, suggesting that different processes are affecting streamwater chemistry at different scales and with different spatial relationships. Our results suggest that streamwater chemistry is more complex than the expected pattern of a single scale of spatial autocorrelation (Figure 1c) or increasing homogeneity with downstream distance. There is both fine-scale patchiness and broad-scale trend operating within the stream network and across the landscape. These observations may provide clues about how streams function and ultimately lead to better models of aquatic and terrestrial ecosystem interactions in stream networks. Furthermore, the spatial structure of stream network chemistry that emerges from our analysis is similar to patterns that have been observed in landscape ecology, where both broad-scale gradients and fine-scale patchiness are influenced by environmental attributes (33).

Semivariograms of water chemistry in the stream network of the Hubbard Brook Valley revealed spatial structure at multiple scales previously described only in unbranched stream sections (11) or for a limited suite of chemical constituents examined at much coarser scales (34). Other pioneering studies of stream networks have examined spatial heterogeneity or developed predictive models in stream networks (16, 35), but these studies have been too coarse in grain (low resolution) to detect spatial structure at scales ranging from hundreds to thousands of meters. Our analysis of the shape and characteristics of semivariograms (Figure 1) of water chemistry in a stream network provide empirical support for three models of spatial structure in stream networks (sensu 26): (1) broad-scale heterogeneity with few patches (Figure 1b; e.g., the Euclidean relationship of Na⁺, dissolved Si, and DIC), (2) fine-scale heterogeneity with many patches (Figure 1c; e.g., flow-connected K⁺, Cl⁻, DOC), and (3) nested heterogeneity that
contains fine-scale patchiness and broad-scale heterogeneity (Figure 1d; e.g., flow-unconnected H\(^+\), Al\(_{soil}\), DOC).

The different spatial relationships permit the examination of heterogeneity dominated by landscape versus stream network processes. Euclidean and flow-unconnected network relationships were associated with the nested type of heterogeneity indicative of landscape influences occurring over multiple spatial scales, such as soil, geology, and vegetation controls on the chemistry of water sources. In contrast, flow-connected relationships were entirely (with the exception of Na\(^+\)) of the single-scale type, suggesting that similarity at large stream distances in the network was largely controlled by channel-mediated transport and groundwater flow accumulation, which seemed to stabilize the variance of all chemical constituents at about 2000 m separation distance. This range among the semivariograms for flow-connected relationships provides an indication of where the hydrologic control on variability begins to overwhelm the patchiness at the fine-scale (<2000 m). As one would expect, hydrologic transport and longitudinal gains of groundwater were the main drivers of the spatial pattern among most of the solutes. At Hubbard Brook, solute chemistry is relatively constant temporally compared to discharge, which may suggest general persistence in these spatial relationships (17).

Euclidean and flow-unconnected semivariograms exhibit higher variance, spatial dependence at broader scales, and multiple structures (single and nested) compared to flow-connected semivariograms, suggesting processes other than hydrologic transport were also responsible for observed patterns of streamwater chemistry. Chemical constituents typically considered to be controlled by mineral weathering (e.g., Ca\(^{2+}\), dissolved Si, DIC, Mg\(^{2+}\), ANC, and Na\(^+\)) all show either broad- or multi-scale heterogeneity with the Euclidean relationship and high nugget values with the flow-unconnected network relationship, suggesting discontinuity or dissimilarity among short unconnected distances such as two sites above a confluence residing on different tributaries (Figure S2). The broad-scale component of variation in the Euclidean relationships may be explained by either gross changes in bedrock type or glacial till composition/thickness across the Hubbard Brook Valley. The Valley is underlain by granodiorite in the western portion and pelitic schist in the eastern portion (36) with an overlying soil parent material composition that varies with lithologic sources eroded and deposited by the most recent glaciation (37). The most apparent broad-scale pattern occurs with Na\(^+\) where the Euclidean semivariogram shows a monotonic increase, which manifests itself spatially with a general increase in concentration across the Valley from west to east. This pattern is different from chloride suggesting the trend in Na\(^+\) is not due to road-salt contamination in the east part of the basin, which is near an interstate roadway. Sodium may be geologically controlled; however, the trend is counter to currently known mineral chemistry differences in the Hubbard Brook Valley (37). Nevertheless, the comparison between the Euclidean and the flow-connected semivariograms for Na\(^+\) suggests a strong landscape control on the spatial pattern at the basin-wide scale, and fine-scale patchiness along the stream. Similar spatial dependence at large Euclidean distance is apparent with K\(^+\), Ca\(^{2+}\), dissolved Si, Cl\(^-\), SO\(_4^{2-}\), ANC, DIC, and specific conductance.

The difference between nugget values of mineral weathering-controlled chemistry (i.e., Ca\(^{2+}\), dissolved Si, Mg\(^{2+}\), ANC, and Na\(^+\)) for the two network relationships suggests differences in landscape character and their potential influence on spatial structure, but over shorter distances. A higher nugget value for the flow-unconnected metric highlights spatial variability over short
distances at tributary junctions and chemical dissimilarities between tributaries. This pattern shows that sources contributing to streamflow that are not downstream from one another may be heterogeneous despite close proximity. This same pattern was noted by Likens and Buso (18) in streams strongly influenced by groundwater seeps and springs. Seeps have been noted throughout the valley and typically show distinct chemistry (e.g., 17).

Streamwater chemistry that exhibited the weakest structure, including NH$_4^+$, NO$_3^-$, and PO$_4^{3-}$, may be explained in large part by their low concentrations (i.e., at or near analytical detection limits [30]). However, there was some indication of spatial dependence over fine scales (<1000 m) in NO$_3^-$ and PO$_4^{3-}$ (i.e., linear increases in variation over short flow-connected distances), suggesting that in-stream uptake may be occurring over very short distances due to biological demand for these nutrients in the generally nutrient-poor environment (38). Uptake lengths for ammonium and phosphorus within the Hubbard Brook Valley were shown to vary between 5 to 277 m and 2 to 54 m, respectively (39), suggesting strong in-stream processing influence on spatial variation over short flow-connected distances.

Both landscape and hydrologic drivers influence some chemical constituents. For example, DOC, H$^+$ and Al$_m$ show autocorrelation at flow-connected relationships >1500 m, indicating an influence of flow accumulation on spatial structure. However, these solutes also exhibited nested structure in flow-unconnected distance at fine (<750 m) and broad (>4500 m) spatial scales. Thus, two scales of heterogeneity and three different spatial relationships appear to be important. One is fine-scale, resulting in hotspots of DOC and H$^+$ variation, which typically coincides with Al$_m$ variation (40). At large distances that are not flow-connected, such as headwater regions in different parts of the network, there is also a high degree of variation in H$^+$, DOC and Al$_m$. This pattern may occur because headwater and valley bottom regions throughout the Hubbard Brook Valley contain patches of coniferous forest and distinct soils. Many of these areas (e.g., ridge tops, north-facing slopes, and along the stream) tend to have shallow, wet, acidic soils (41) with deep litter layers, and high DOC and Al soil-water concentrations that have been shown to affect spatial patterns of streamwater DOC and Al (17, 40). It is striking that our semivariogram analysis provides a quantitative characterization of this large-scale patchiness, which small-scale, descriptive studies cannot capture.

This study shows that network geostatistics combined with high-resolution water chemistry data can provide insight into dominant processes driving biogeochemical patterns in stream networks at local, longitudinal, and landscape scales. In the Hubbard Brook Valley, landscape processes were shown to occur over a wide range of spatial scales (500 to >6000 m), with broad-scale trends and fine-scale patchiness likely driven by geologic, soil, and vegetation features across the Valley. Hydrologic transport and flow accumulation in streams were also found to exert influence on streamwater chemistry at downstream distances of 1400 to 2200 m, where variation in chemistry tended to stabilize. By revealing the spatial structure of stream biogeochemistry and scales of underlying drivers, this spatially explicit, network-level analysis is crucial to refining long-held assumptions about stream structure and function (e.g., 42).

**Methods**

*Basin-wide survey*
Streamwater chemistry was measured by Likens and Buso (18) throughout the 3600 ha Hubbard Brook Valley in the White Mountains of central New Hampshire (USA) (43°56′N, 71°45′W). Detailed information on the ecological, hydrological, climatological, and geological setting is summarized by Likens (43). The field survey of streamwater chemistry (Al₃m, acid-neutralizing capacity, Ca²⁺, Cl⁻, DIC, DOC, K⁺, Mg²⁺, Na⁺, NH₄⁺, NO₃⁻, pH, PO₄³⁻, dissolved Si, SO₄²⁻, and specific conductance) was conducted during October through December 2001 and is described in Likens and Buso (18). A subset of 664 from the total of 761 samples was selected for network analysis based on an assessment of hydrography and the ability to delineate catchment areas using a geographical information system (GIS) derived from 10-m digital elevation models. The sites left out were almost entirely small tributaries with only a few samples and they were scattered systematically throughout the Valley, but primarily in headwater areas.

**Analysis of spatial structure with empirical semivariograms**

We quantified spatial structure in streamwater chemistry by calculating empirical semivariograms using distance metrics based on spatial connectivity within the stream network and across the landscape of the Hubbard Brook Valley (20, 21). Empirical semivariograms provide a means to describe spatial structure in geographically referenced data by quantifying the variance (or relatedness) in sample values (e.g., Ca²⁺ or specific conductance) as a function of the distance between sample points in Euclidean space (44) or throughout the stream network (24, 45). An empirical semivariogram that displays semivariance as a function of stream network distance separately for flow-connected and flow-unconnected relationships is called a Torgegram (46). We used Torgegrams and empirical semivariograms based on Euclidean distance as exploratory tools for visualizing patterns of spatial autocorrelation.

Empirical semivariograms provide information on patterns of spatial dependence and have various theoretical forms (Figure 1). We used the typology proposed by Ettema and Wardle (26) to compare shapes of semivariograms among chemical constituents and spatial relationships. Because all of the spatial relationships used the same data, standardization of semivariance was not necessary for comparisons among relationships for a given chemical constituent. Only the shapes of the semivariograms were compared among chemical constituents, and this approach made it possible to preserve the original units in the semivariograms to aid interpretation. In geostatistics, terms used to describe the semivariogram include the range, sill, and nugget, all of which can be estimated quantitatively and have specific meaning for the purposes of predictive modeling (47). We did not attempt fit models to the empirical semivariograms because of the complexity of the structures observed and uncertainty associated with identifying model forms a priori to estimate parameters. We determined the approximate range visually by identifying the inflection point (or points) in variance as a function of separation distance. The approximate variances at the y-intercept (i.e., the nugget) and the inflection point in variance (i.e., the sill) were also assessed visually. The goal of analysis with semivariograms was to provide a means to visualize, with the aid of relatively simple geostatistical tools, the full range of spatial complexity in biogeochemical patterns across scales in a headwater stream network. Specific elucidation of mechanisms and explanatory variables for individual chemical constituents is the topic of ongoing work, including the examination of chemical variability at scales finer than that available from Likens and Buso (18) (see 17).

**Acknowledgements**
Financial support data collection was provided by the National Science Foundation (NSF) including the LTREB and LTER programs, and The A.W. Mellon Foundation. Support for KJM and SWB was provided by NSF EAR 1014507. Support for WHL was provided by NSF DEB 1050459. The U.S. Forest Service (NRS, Newtown Square, PA) operates and maintains the HBEF. This manuscript benefited greatly from discussions with the Spatial Statistics for Streams Working Group supported by NCEAS (NSF EF 0553768), UC Santa Barbara, and the State of California. E. Peterson and N. Grimm provided critical reviews that greatly improved earlier versions of this manuscript. P. Likens is thanked for help with submission of the manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Figure 1. Hypothetical semivariograms and associated maps depicting representative spatial patterns of water chemistry in a stream network. Nonstructured spatial pattern (a) is indicated in the semivariogram by no change in semivariance ($\gamma$) ($y$-axis) with increasing distance ($d$) between neighbors ($x$-axis), as is graphically depicted by the uniform line color in the associated network map. In the example shown (a), $\gamma = 0$ for a uniform, nonstructured spatial pattern. Other potential semivariograms and associated network patterns include (b) spatial dependence at a broad-scale with a gradient symbolized in the network map by changes in line color from the upper left (blue) to the lower right (red) of the stream network, (c) fine-scale patchiness or spatial dependence indicated in the network map as ‘hotspots’, and (d) nested heterogeneity reflecting a combination of fine-scale patchiness imbedded within a broad-scale gradient (sensu Ettema and Wardle 2002). Characteristics of the semivariogram (c) are the asymptote or sill, which is roughly equivalent to the total population variance; the variance discontinuity at the $y$-intercept or nugget, which represents variance due to sampling error and/or spatial dependence at distance intervals not explicitly sampled; and the range, which defines the distance or scale over which spatial dependence is expressed. Beyond this range, in a non-nested structure, points are spatially independent of one another or uncorrelated. Nested semivariograms are hierarchical structures, each characterized by its own range.

Figure 2. Spatial patterns of (a) sodium ($\text{Na}^+$; mg/L), (b) dissolved silica (DSi; reported as SiO$_2$ mg/L), (c) dissolved organic carbon (DOC; mg/L), and (d) specific conductance (SC; $\mu$S/cm) derived from sampling streamwater chemistry at 664 locations throughout the Hubbard Brook Valley, NH during October–December 2001.

Figure 3. Empirical semivariograms for (a) sodium ($\text{Na}^+$; mg/L), (b) dissolved silica (DSi; reported as SiO$_2$ mg/L), (c) dissolved organic carbon (DOC; mg/L), and (d) specific conductance (SC; $\mu$S/cm) based on Euclidian, flow-connected, and flow-unconnected spatial relationships ($d$) in the Hubbard Brook Valley. Visually estimated the ranges (see Table 1) for Euclidean (E) (yellow and black line), flow-connected (FC) (orange), and flow-unconnected (FU) (blue) relationships are indicated with vertical lines. Symbols with lighter color shades indicate semivariance estimates based on < 100 pairs of points.
Characteristics of empirical semivariograms for streamwater chemistry based on Euclidean, flow-connected, and flow-unconnected spatial relationships in the Hubbard Brook Valley. Semivariogram type and range for each spatial relationship indicate nested- versus single-scale spatial structure and the distances over which spatial dependence is expressed, respectively.

<table>
<thead>
<tr>
<th>Solute</th>
<th>Euclidean relationship</th>
<th>Flow-connected network relationship</th>
<th>Flow-unconnected network relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td>Scale (m)</td>
<td>Type</td>
</tr>
<tr>
<td>H⁺</td>
<td>single</td>
<td>2300</td>
<td>single</td>
</tr>
<tr>
<td>Ca²⁺</td>
<td>nested</td>
<td>1000</td>
<td>single</td>
</tr>
<tr>
<td>Mg²⁺</td>
<td>nested</td>
<td>1100</td>
<td>single</td>
</tr>
<tr>
<td>Na⁺</td>
<td>single</td>
<td>?</td>
<td>nested</td>
</tr>
<tr>
<td>K⁺</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>Al₃⁺</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>DSO₄²⁻</td>
<td>nested</td>
<td>200</td>
<td>single</td>
</tr>
<tr>
<td>Cl⁻</td>
<td>nested</td>
<td>200</td>
<td>single</td>
</tr>
<tr>
<td>NO₃⁻</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>PO₄³⁻</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>DOC</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>DIC</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>ANC</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
<tr>
<td>SC</td>
<td>nested</td>
<td>1200</td>
<td>single</td>
</tr>
</tbody>
</table>

The dash symbol (-) indicates semivariograms (1) for which the type or range could not be determined due to low chemical concentrations or (2) that are not applicable for the given type. An unknown range is indicated with a question mark (?) for solutes with semivariograms that appear to be nested beyond the maximum separation distance used for the analysis or for ranges that cannot easily be determined visually.
Distance ($d$)

Range

Sill

Nugget

Small scale

Large scale

Distance ($d$)
Supplemental Methods

Basin-wide survey

The study area is located largely within the Hubbard Brook Experimental Forest (HBEF) where biogeochemistry and ecosystem processes have been investigated by the HBES for almost five decades (1). The watershed is vegetated with hardwood and mixed hardwood–conifer forest interspersed with conifer patches. Elevations in the watershed range from 200 m at the outlet to over 1000 m in the headwaters. Detailed information on the ecological, hydrological, climatological, and geological setting is summarized by Likens and Buso (2), Likens (1) and the HBES (www.hubbardbrook.org). Water samples for this study were collected along the main stem of Hubbard Brook at locations directly upstream of tributary junctions (17 sites along the main stem), whereas samples in the tributaries were collected at 100-m intervals throughout 32 tributaries across the entire fifth-order stream network (total length surveyed = 75 km) (see Likens and Buso [2006] for details on field and analytical methods). Although the density of samples was much higher in the tributaries than in the main stem, longitudinal patterns in the main stem revealed relatively homogeneous solute concentrations and justified a coarser sampling density.

Analysis of spatial structure with empirical semivariograms

Matrices of distances between pairs of points used in the calculation of semivariograms were calculated in a GIS based on Euclidean distance and on stream network distance with software specifically designed for analyzing stream networks (3). The GIS software for stream network analysis created a matrix of downstream-only distances that was manipulated to produce matrices of stream network distances based on flow-connected and flow-unconnected relationships (4). Semivariograms were calculated with customized functions in the S-PLUS statistical package (5, 6).

The classical estimator of the semivariogram is sensitive to outliers and non-normal distributions inherent in ecological data. Therefore, we used the robust estimator of semivariance recommended by Cressie (7). Broad- and fine-scale spatial trends are typically removed from the data by detrending before semivariograms are calculated (8). This approach is particularly important when the semivariograms are used for prediction at unsampled locations or when broad-scale trends are known and, thus, not the focus of analysis (9) (e.g., the effects spatial streamflow accumulation). We investigated potential trends in streamwater chemistry with respect to catchment area as a surrogate for streamflow accumulation and downstream distance (cf. 10) and found only weak relationships with these variables (r < |0.13| and only 6 out of 16 variables had p<0.05). Moreover, removing these trends had little or no effect on the shape or characteristics of the empirical semivariograms. Therefore, we used the untransformed data in our analyses.

The number of pairs of points for a given distance class in the semivariogram varied as a function of distance and among the various spatial relationships, but was consistent among streamwater chemistry variables. Because low numbers of pairs for a given distance class can complicate the interpretation of the semivariogram, we verified that the number of pairs for each semivariance calculation exceeded the Rossi et al. (11) recommendation of 50 pairs per distance class; the numbers of pairs for flow-connected, flow-unconnected, and Euclidean spatial relationships were 78-679, 62-3433, and 569-3527, respectively. Semivariance was calculated at
100-m intervals with a maximum separation distance between any two points in the stream network of 6950 m. At the maximum separation distance, at least 100 distance pairs were included for all relationships (106, 2396, and 642 for flow-connected, flow-unconnected, and Euclidean relationships, respectively).

References
Legends for Supplemental Figures

Figure S1 (Supplement). Spatial patterns of (a) hydrogen ion (eq/L), (b) acid-neutralizing capacity (ANC; µeq/L), (c) calcium (Ca\(^{2+}\); mg/L), (d) magnesium (Mg\(^{2+}\); mg/L), (e) potassium (K\(^{+}\); mg/L), (f) sulfate (SO\(_4^{2-}\); mg/L), (g) ammonium (NH\(_4^{+}\); mg/L), (h) nitrate (NO\(_3^{-}\); mg/L), (i) ortho-phosphate (PO\(_4^{3-}\); mg/L), (j) chloride (Cl\(^{-}\); mg/L), (k) total monomeric aluminum (Al\(_m\); mg/L), and (l) dissolved inorganic carbon (DIC; µmol/L).

Figure S2 (Supplement). Empirical semivariograms for (a) hydrogen ion (eq/L), (b) calcium (Ca\(^{2+}\); mg/L), (c) magnesium (Mg\(^{2+}\); mg/L), (d) potassium (K\(^{+}\); mg/L), (e) ammonium (NH\(_4^{+}\); mg/L), (f) total monomeric aluminum (Al\(_m\); mg/L), (g) sulfate (SO\(_4^{2-}\); mg/L), (h) chloride (Cl\(^{-}\); mg/L), (i) nitrate (NO\(_3^{-}\); mg/L), (j) ortho-phosphate (PO\(_4^{3-}\); mg/L), (k) dissolved inorganic carbon (DIC; µmol/L), and (l) acid-neutralizing capacity (ANC; µeq/L) based on Euclidian, flow-connected, and flow-unconnected spatial relationships (d) in the Hubbard Brook Valley. Symbols with lighter color shades indicate semivariance estimates based on < 100 pairs of points.