Supporting Information

McGuire et al. 10.1073/pnas.1404820111

SI 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 Hubbard Brook Ecosystem Study (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 more than 1,000 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 (2) 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 geographical information system (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 nonnormal distributions inherent in ecological data. Therefore, we used the robust estimator of semivariance recommended by Cressie (7). Broad-scale 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 (compare ref. 10) and found only weak relationships with these variables ($r < |0.13|$, and only 6 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, 62–3,433, and 569–3,527, respectively. Semivariance was calculated at 100-m intervals with a maximum separation distance between any two points in the stream network of 6,950 m. At the maximum separation distance, at least 100 distance pairs were included for all relationships (106, 2,396, and 642 for flow-connected, flow-unconnected, and Euclidean relationships, respectively).

1. Likens GE (2013) *Biogeochemistry of a Forested Ecosystem* (Springer, New York), 3rd Ed, p 208.
2. Likens GE, Buso DC (2006) Variation in streamwater chemistry throughout the Hubbard Brook Valley. *Biogeochemistry* 78:1–30.
3. Theobald DM, et al. (2006) *Functional Linkage of Water Basins and Streams (FLoWS): ArcGIS tools for Network-Based Analysis of Freshwater Ecosystems* (Nat Resour Ecol Lab, Colorado State Univ, Fort Collins, CO), Version 1.
4. Peterson EE, Theobald DM, Ver Hoef JM (2007) Geostatistical modelling on stream networks: Developing valid covariance matrices based on hydrologic distance and stream flow. *Freshw Biol* 52:267–279.
5. Ganio LM, Torgersen CE, Gresswell RE (2005) A geostatistical approach for describing spatial pattern in stream networks. *Front Ecol Environ* 3(3):138–144.
6. S-PLUS (2002) *S-PLUS for Windows, Version 6.1* (Insightful Corp, Seattle).
7. Cressie NAC (1993) *Statistics for Spatial Data* (John Wiley & Sons, New York).
8. Chiles J-P, Delfiner P (1999) *Geostatistics: Modeling Spatial Uncertainty* (John Wiley & Sons, New York).
9. Journel AG, Rossi ME (1989) When do we need a trend model in kriging? *Math Geol* 21(7):715–739.
10. Ver Hoef J, Peterson E, Theobald DM (2006) Spatial statistical models that use flow and stream distance. *Environ Ecol Stat* 13:449–464.
11. Rossi RE, Mulla DJ, Journel AG, Franz EH (1992) Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecol Monogr* 62:277–314.

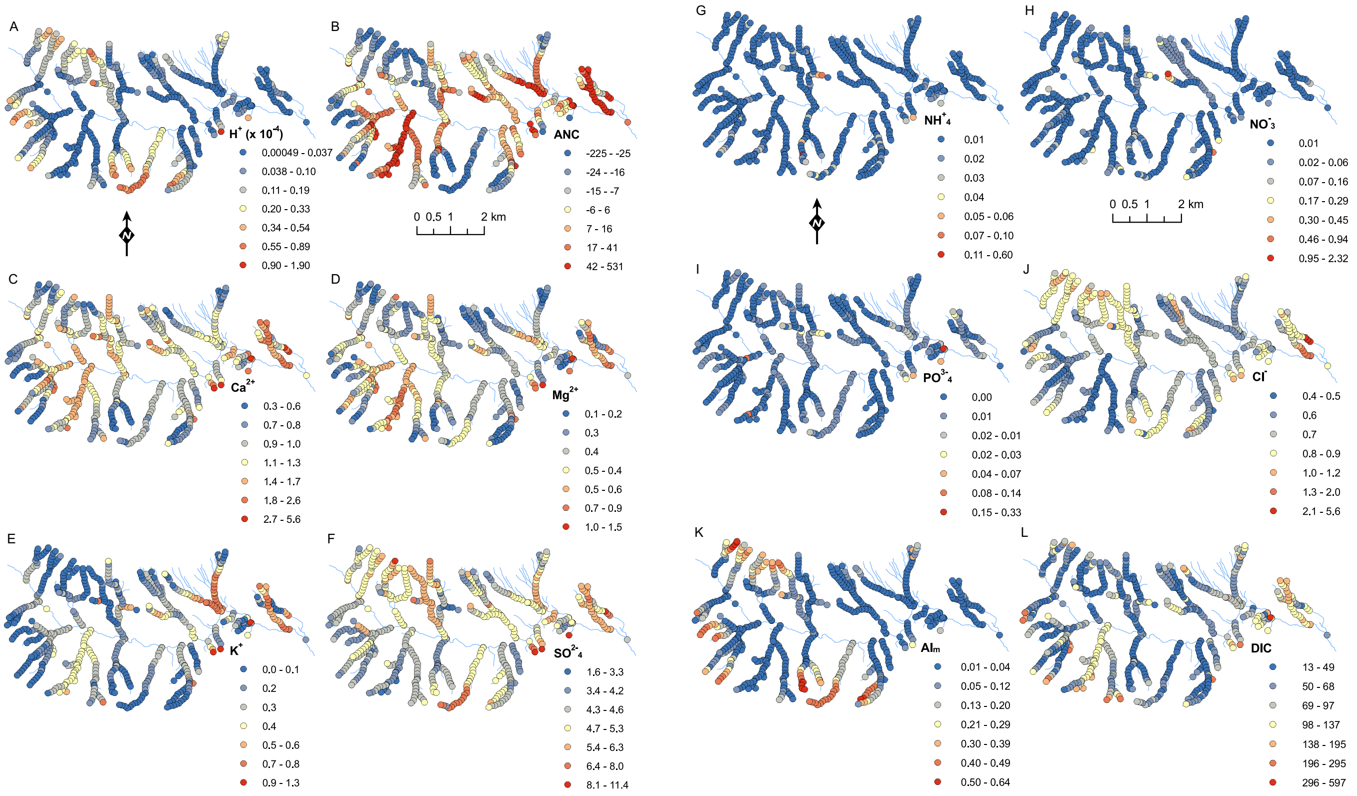

Fig. S1. Spatial patterns of hydrogen ion (eq/L) (*A*), acid-neutralizing capacity (ANC; μeq/L) (*B*), calcium (Ca^{2+}; mg/L) (*C*), magnesium (Mg^{2+}; mg/L) (*D*), potassium (K^+; mg/L) (*E*), sulfate (SO_4^{2-}; mg/L) (*F*), ammonium (NH_4^+; mg/L) (*G*), nitrate (NO_3^-; mg/L) (*H*), ortho-phosphate (PO_4^{3-}; mg/L) (*I*), chloride (Cl^-; mg/L) (*J*), total monomeric aluminum (Al_m; mg/L) (*K*), and dissolved inorganic carbon (DIC; μmol/L) (*L*).

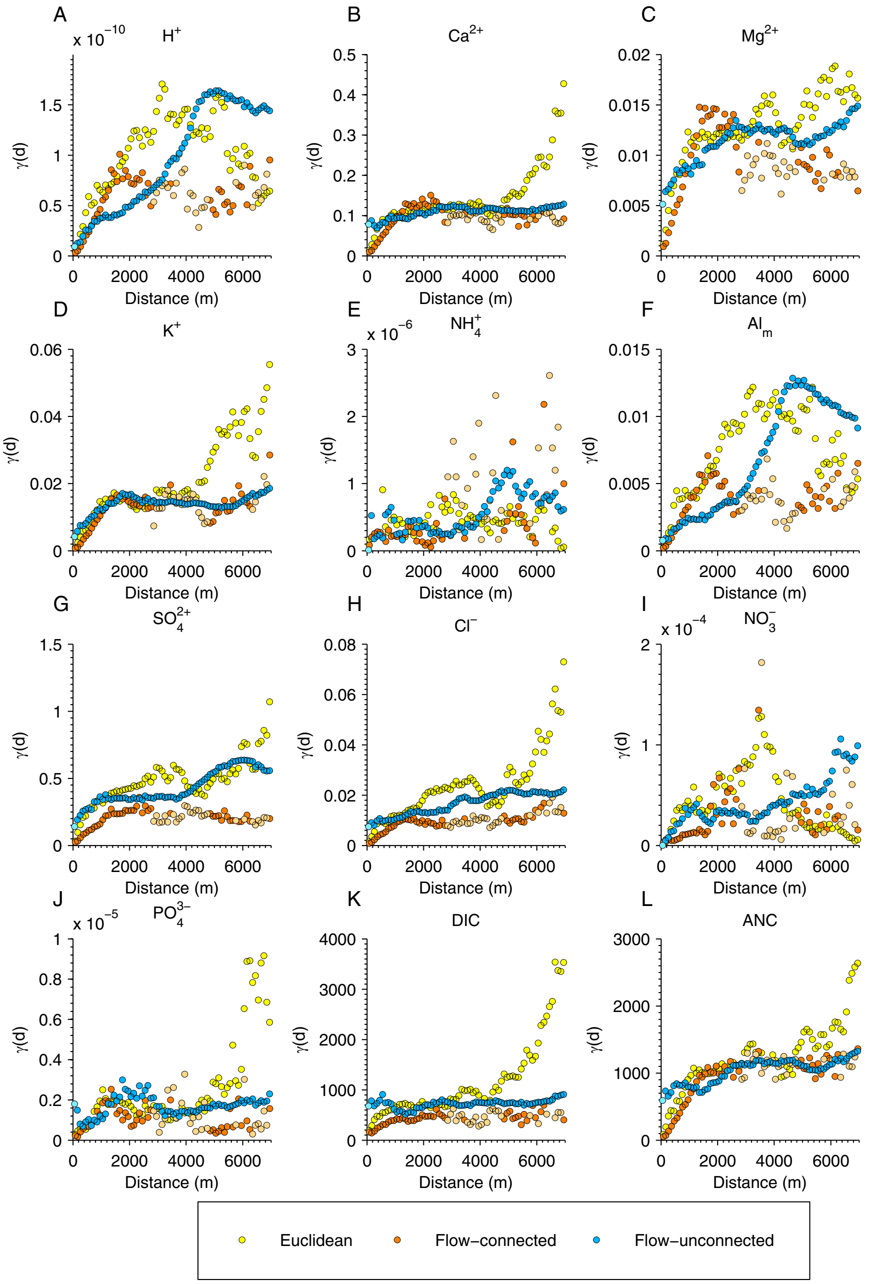

Fig. S2. Empirical semivariograms for hydrogen ion (eq/L) (*A*), calcium (Ca^{2+}; mg/L) (*B*), magnesium (Mg^{2+}; mg/L) (*C*), potassium (K^{+}; mg/L) (*D*), ammonium (NH_4^{+}; mg/L) (*E*), total monomeric aluminum (Al_m; mg/L) (*F*), sulfate (SO_4^{2-}; mg/L) (*G*), chloride (Cl^{-}; mg/L) (*H*), nitrate (NO_3^{-}; mg/L) (*I*), ortho-phosphate (PO_4^{3-}; mg/L) (*J*), dissolved inorganic carbon (DIC; μmol/L) (*K*), and acid-neutralizing capacity (ANC; μeq/L) (*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.