Supplement of Hydrol. Earth Syst. Sci., 25, 3397–3409, 2021
https://doi.org/10.5194/hess-25-3397-2021-supplement

Supplement of

Machine learning deciphers CO_2 sequestration and subsurface flowpaths from stream chemistry

Andrew R. Shaughnessy et al.

Correspondence to: Andrew R. Shaughnessy (ars637@psu.edu)

Section 1 NMF Model

To employ NMF on limited datasets of stream chemistry, a bootstrapped data set was generated using a multivariate normal distribution of log-transformed stream water chemistries, similar to the procedure outlined in Lautz et al. (2014). The bootstrapped dataset matches the measured means of the log-transformed stream water chemistries and maintains covariation between analytes. A comparison between the measured and bootstrapped data sets can be seen in Fig. S2. All of the input features were normalized to values between 0 and 1 to not bias the model training to any one input feature. Next, the model was trained to the bootstrapped dataset using NMF algorithms in the python library scikit-learn (Pedregosa et al., 2011). Lastly, the trained model was applied to measured stream water samples to delineate mixing proportions.

The model results are sensitive to the random initiation of the *H* matrix (i.e., endmember chemistries) used in the training. To produce a more robust decomposition, the starting *H* matrix was randomly initiated 20,000 times. For each stochastic iteration, we used NMF to calculate optimal *W* and *H* matrices (Eq 1) and then filtered out any models with proportions that did not add to 1 ± 0.05. Additionally, the fit of the model was evaluated from SSE:

$$SSE = \sum_m \left(\left(\frac{[X_m]}{[SO_4^{2-}]} \right)_n - \sum_p \alpha_p \left(\frac{[X_m]}{[SO_4^{2-}]} \right)_p \right)^2, \tag{S1}$$

Here, SSE is the sum of square errors, α_p is the sulfate mixing proportion of endmember *p* derived from the model, *X* is the element "m" (i.e., Ca, Mg, Na, K, and Cl), and brackets denote concentration. The subscript "*n*" refers to measured concentrations at timepoint "*n*" in the stream. Using eq. S1, we filtered out additional models that yielded poor fitting solutions following the procedure outlined in Torres et al. (2016). Here we define a reference SSE that is equal to the 5th percentile SSE for all the models for that sample and filtered out any models where the SSE was larger than the reference SSE. In other words, we kept only the 5th percentile of best fitting models for each sample. The remaining models were averaged and reported as the final result. Additionally, we calculate the standard deviation for the remaining samples to represent uncertainty in our modeling results and we propagate these errors throughout our calculations. This average number of valid models per sample after all filters were applied was 44 for Shale Hills, 104 for East River, and 55 for Hubbard Brook.

Section 2 Calculations

S2.1 Solute Fluxes

The time-averaged flux of each species, *Flux*, was calculated using values for concentration and discharge following an equation adapted from Moatar et al. (2013):

$$Flux = \beta \frac{\sum_n [X]_n Q_n}{\sum_n Q_n} \bar{q} A^{-1}, \tag{S2}$$

Here, *Flux* has the units of mmol m^{-2} yr^{-1}, $[X]_n$ is the concentration of a weathering product in the stream (e.g., X = SO_4^{2-}, Ca^{2+}, Mg^{2+}, Na^+, K^+, Cl^-) at timepoint n, Q_n is the discharge measured at timepoint n, $\bar{q}$ is the mean discharge over the period of record (i.e., sum of daily discharge measurements/number of daily discharge measurements), A is the basin area (Shale Hills = 8.0×10^4 m^2; East River = 8.5×10^7 m^2; Hubbard Brook = 4.0×10^5 m^2 (W3), 1.5×10^5 m^2 (W6), 7.7×10^5 m^2 (W7), 6.1×10^5 m^2 (W8), 7.0×10^5 m^2 (W9)), and β is a coefficient for unit conversions to mmol m^{-2} yr^{-1} (i.e., 3.15×10^{10} l s m^{-3} yr^{-1}) or to meq m^{-2} yr^{-1} (i.e., 3.15×10^{10} l s m^{-3} yr^{-1} multiplied by species charge (meq/mmol)).

We calculate the uncertainty for our flux calculations by propagating errors in $[X]_n$ from equation S2. For sulfate concentrations, we assume 5% error on the measured concentrations. Uncertainties in mixing proportions (i.e., α) from the NMF model are utilized as uncertainties in the pyrite-derived and acid rain-derived sulfate concentrations. Lastly, error in the total cation concentration is determined from the uncertainties in the NMF modeled total cation concentrations. For Hubbard Brook, there are 5 sub-catchments that we use in analysis. Reported fluxes in Table 1 are the average of the sub-catchments.

S2.2 Using Stream Chemistry to Calculate CO_2 Drawdown or Release

Here we calculate the inferred CO_2 release or sequestration resulting from weathering as recorded in the sum of all base cation concentrations (meq/l) in each stream sample, $[\Sigma^+]_{total}$:

$$[\Sigma^+]_{total} = 2[Ca^{2+}]_{total} + 2[Mg^{2+}]_{total} + [Na^+]_{total} + [K^+]_{total}, \tag{S3}$$

Here, we use the modeled base cation concentrations from NMF in eq. S3, and we use the uncertainty in the modeled concentrations for the error in $[\Sigma^+]_{total}$. To calculate the inferred CO_2 release or sequestration resulting from weathering, we use the results of NMF, as described below, to identify the extents of 4 weathering reactions recorded in each stream sample: 1) CO_2-driven weathering (CO_2-weathering) of silicates, 2) H_2SO_4-driven weathering (H_2SO_4-weathering) of silicates, 3) CO_2-weathering of carbonates, and 4) H_2SO_4-weathering of carbonates. We note these four quantities respectively as 1) $[\Sigma^+]_{carbonate-CO_2}$; 2) $[\Sigma^+]_{silicate-H_2SO_4}$; 3)$[\Sigma^+]_{silicate-CO_2}$; 4)$[\Sigma^+]_{carbonate-H_2SO_4}$. These are the four unknowns we seek to calculate for SH and ER, as described below.

Based on the high proton and low metal concentrations of the measured rain chemistry, the rain contributes negligibly to the base cation concentrations of the study streams; therefore, we apportioned all the base cations to weathering reactions. First, we note that the meq/l of cations derived from carbonate minerals, $[\Sigma^+]_{carbonate}$, equal $[\Sigma^+]_{carbonate-CO_2} + [\Sigma^+]_{carbonate-H_2SO_4}$. Likewise, the meq/l of cations derived from silicate minerals, $[\Sigma^+]_{silicates}$ equal $[\Sigma^+]_{silicate-H_2SO_4} + [\Sigma^+]_{silicate-CO_2}$. The summation of silicate-cations ($[\Sigma^+]_{silicate}$) is the difference between the summation of total cations ($[\Sigma^+]_{total}$) and that of carbonate-derived cations ($[\Sigma^+]_{carbonate}$):

$$[\Sigma^+]_{silicate} = [\Sigma^+]_{total} - [\Sigma^+]_{carbonate}, \tag{S4}$$

We use a few field observations to complete the calculations for SH and ER, as explained in the main text. First, carbonate minerals only dissolve in water flowing along the deep path because carbonates have been depleted from shallow depths. Second, although some chlorite dissolves into water flowing along the deep path, the release of Mg at depth is insignificant compared to Mg released from carbonate. So we ignore shallow dissolution of carbonates and deep dissolution of silicates in both SH and ER.

With these observations, we can write:

$$[\Sigma^+]_{carbonate} = 2\alpha_{deep}[SO_4^{2-}]_{total}\left(\left(\frac{[Ca^{2+}]}{[SO_4^{2-}]}\right)_{deep} + \left(\frac{[Mg^{2+}]}{[SO_4^{2-}]}\right)_{deep}\right), \tag{S5}$$

Here, α_{deep} is the proportion determined through NMF of sulfate in a given water sample that was derived from reactions along the deep flowpath, $[SO_4^{2-}]_{total}$ is the total concentration of sulfate in the stream water sample under consideration, $([Ca^{2+}]/[SO_4^{2-}])_{deep}$ and $([Mg^{2+}]/[SO_4^{2-}])_{deep}$ are the model-derived ratios of $[Ca^{2+}]$ and $[Mg^{2+}]$ to $[SO_4^{2-}]$, respectively, that characterize the deep flowpath endmember for that sample.

Remembering that Mg release from chlorite dissolution at depth is insignificant compared to Mg from carbonates, all of the generated sulfate in the deep weathering endmember is balanced by cations from dissolved carbonate minerals:

$$[\Sigma^+]_{carbonate-H_2SO_4} = 4\alpha_{deep}[SO_4^{2-}]_{total}, \tag{S6}$$

(We multiply the concentration of deep sulfate by 4 because 4 eq of cations are released per mol of sulfate, noting that $[\Sigma^+]$is in eq/L and $[SO_4^{2-}]$ is in mol/L). Any carbonate-derived base cations that are in excess of what could have been produced by pyrite-derived sulfuric acid are attributed to CO_2-weathering of carbonates:

$$[\Sigma^+]_{carbonate-CO_2} = [\Sigma^+]_{carbonate} - [\Sigma^+]_{carbonate-H_2SO_4}, \tag{S7}$$

Remembering that no carbonates dissolve into water flowing along the shallow path, then similar arguments for the shallow flowpath yield:

$$[\Sigma^+]_{silicate-H_2SO_4} = 2\alpha_{shallow}[SO_4^{2-}]_{total}, \tag{S8}$$

$$[\Sigma^+]_{silicate-CO_2} = [\Sigma^+]_{silicate} - [\Sigma^+]_{silicate-H_2SO_4}, \tag{S9}$$

From these equations, values for the four unknowns can be calculated for SH and ER. A similar approach was taken for HB except that no carbonate minerals were present, and only two unknowns were determined ($[\Sigma^+]_{silicate-H_2SO_4}$, $[\Sigma^+]_{silicate-CO_2}$).

With respect to the atmosphere considered over the long-term (10^5-10^6 yr), H_2SO_4-weathering of silicates and CO_2-weathering of carbonates are CO_2 neutral, while CO_2-weathering of silicates sequesters CO_2 and H_2SO_4-weathering of carbonates releases CO_2 (Fig. 1). As seen in Figure 1, per mole of $CaSiO_3$ or $CaCO_3$ weathered, CO_2-weathering of silicates sequesters 1 mol of CO_2 and H_2SO_4-weathering of carbonates releases 0.5 moles of CO_2. In terms of $[\Sigma^+]_{total}$, CO_2-weathering of silicates sequesters 0.5 moles of CO_2 per base cation equivalent released into solution and H_2SO_4-weathering of carbonates releases 0.25 moles of CO_2 per base cation equivalent released into solution (Fig. 1; Reactions 2, 3, 6, and 7). For a given water sample, the cation concentrations record the extent of dissolution of carbonate and silicates, as long as the contribution of these base cations from acid rain is minimal. (For simplicity, we do not correct $[\Sigma^+]$ for rain chemistry but see SM Section 4). Therefore, the uptake or release of CO_2, ΔCO_2, can be calculated for any given stream water sample:

$$\Delta CO_2 = 0.5\,[\Sigma^+]_{silicate-CO_2} - 0.25\,[\Sigma^+]_{carbonate-H_2SO_4}. \tag{S10}$$

Using ΔCO_2, we calculate the flux of CO_2 using the discharge measurements for each sample (see Fig. S4).

Next, we will derive κ_{stream}, the modern CO_2 sequestration coefficient. In general, both κ_{stream} and κ_{rock} (see SM 2.3) are used as ways to note the extent that weathering in a watershed is sequestering or releasing CO_2. κ_{stream} is the amount of CO_2 emitted or sequestered calculated from $[\Sigma^+]_{total}$ as described above, normalized by $[\Sigma^+]_{total}$(meq/l):

$$\kappa_{stream} = -\frac{\Delta CO_2}{[\Sigma^+]_{total}}, \tag{S11}$$

The negative sign is used so that a negative κ_{stream} represents sequestration (uptake of CO_2), and a positive κ_{stream} represents release. From eq. S11 it is apparent that the CO_2 emitted or sequestered equals the product, κ_{stream} $[\Sigma^+]_{total}$, with the appropriate sign. Total dissolved base cations in a stream draining a watershed with no carbonate nor pyrite are attributed here entirely as CO_2-weathering: this watershed demonstrates the highest capacity to sequester CO_2 and κ_{stream} equals -0.5. Substituting from eq. S10 into eq. S11 yields:

$$\kappa_{stream} = -\frac{0.5\,[\Sigma^+]_{silicate-CO_2} - 0.25\,[\Sigma^+]_{carbonate-H_2SO_4}}{[\Sigma^+]_{total}}, \tag{S12}$$

We can further expand eq. S12 by substituting eq. S6 for $[\Sigma^+]_{\text{carbonate-H2SO4}}$, eq. S9 for $[\Sigma^+]_{\text{silicate-CO2}}$, eq. S4 for $[\Sigma^+]_{\text{silicate}}$ and eq. S8 for $[\Sigma^+]_{\text{silicate-H2SO4}}$

$$\kappa_{stream} = -\frac{0.5\left([\Sigma^+]_{total} - [\Sigma^+]_{carbonate} - 2\alpha_{shallow}[SO_4^{2-}]_{total}\right) - \alpha_{deep}[SO_4^{2-}]_{total}}{[\Sigma^+]_{total}}, \quad \text{(S13)}$$

This can be rearranged and simplified as:

$$\kappa_{stream} = -\frac{1}{2} + \frac{1}{2}\frac{[\Sigma^+]_{carbonate}}{[\Sigma^+]_{total}} + \frac{[SO_4^{2-}]_{total}}{[\Sigma^+]_{total}}, \quad \text{(S14)}$$

We then define the second term (ratio of carbonate-derived base cations to total base cations in the stream sample) as γ_{stream} and the third term (ratio of the sulfate equivalents (from sulfuric acid) to the equivalents of base cations in the stream) as ζ_{stream}. Note that to obtain the sulfate equivalents, we multiply $[SO_4^{2-}]_{\text{total}}$ by 2, resulting in the third term equal to $0.5\zeta_{\text{stream}}$. Given these definitions, eq. S14 yields eq. 2 from the main text:

$$\kappa_{stream} = \frac{1}{2}(-1 + \gamma_{stream} + \zeta_{stream})$$

S2.3 Using Rock Chemistry to Calculate CO_2 Drawdown or Release

Here we compare the bulk elemental composition of parent rock to topsoil and calculate the difference to determine if the system acted on net as a source or a sink of CO_2 over the weathering duration. Of course, this calculation involves inspection only of rock versus soil chemistry and cannot therefore be used to separate CO_2- versus H_2SO_4-weathering when the latter is derived from acid rain. The three most important factors are i) the ratio of base cations in carbonates relative to silicates in the rock, ii) the ratio of acid-generating units of pyrite relative to total base cations in carbonate+silicate minerals, and iii) the ratio of base cations still retained in regolith at the land surface relative to total base cations. This latter ratio is related to the chemical depletion factor (written below as $-\tau$), i.e., the relative ratio of loss of a component in a rock to chemical weathering versus total loss by physical + chemical weathering (Riebe et al. 2003). For (i), we define the carbonate/silicate factor, γ_{rock}, which is the proportion of base cation equivalents in the rock derived from carbonate minerals divided by the total base cations:

$$\gamma_{rock} = \frac{2C_{Ca,carb} + 2C_{Mg,carb}}{2C_{Ca,Total} + 2C_{Mg,Total} + C_{Na,Total} + C_{K,Total}}, \quad \text{(S15)}$$

Here $C_{X,k}$ is the mole fraction (mol/kg) of base cation (X = Ca, Mg, Na, or K) in carbonates (k = carb) or in carbonate + silicate minerals (k = Total). By definition, γ_{rock} ranges from 0 (where all base cations derive from silicates) to 1 (where all base cations derive from carbonates). Likewise, 1- γ_{rock} is the proportion of base cations derived from silicate minerals.

When pyrite oxidizes it produces sulfuric acid that can dissolve carbonate and silicate minerals. This impacts CO_2 dynamics over 10^5-10^6 yr timescales by releasing CO_2 (H_2SO_4-weathering of carbonates). But it also diminishes the silicate content of the rock, thereby diminishing the rock's capacity to sequester CO_2. Here, we define a new variable, ζ_{rock}, which is the acid generation capacity expressed relative to the base cations in the rock (all on an equivalents basis):

$$\zeta_{rock} = \frac{4C_{py}}{2C_{Ca,Total}+2C_{Mg,Total}+C_{Na,Total}+C_{K,Total}}, \tag{S16}$$

Here, the subscript *py* refers to pyrite (mol/kg rock). We multiply the concentration of pyrite (i.e., C_{py}) by 4 (eq/mol) because 4 equivalents of sulfate are produced per mole of pyrite as shown in reaction S17.

$$FeS_2 + \frac{15}{4}O_2 + 2H_2O \rightarrow \frac{1}{2}Fe_2O_3 + 4H^+ + 2SO_4^{2-}, \tag{S17}$$

Lastly, in many catchments, the bulk chemistry of parent rock is not indicative of the CO_2 sequestration during weathering because silicate minerals are kinetically slow to dissolve and they do not completely dissolve before the rock physically erodes. On the other hand, we assume here that all carbonate minerals chemically weather away before exposure at land surface, an assumption most useful for wet climates and relatively low-carbonate content rocks. The relative depletion of an element in a weathered rock with respect to the parent rock is easily calculated from the mass transfer coefficient, τ.

$$\tau_{i,j} = \frac{C_{j,weathered}C_{i,parent}}{C_{j,parent}C_{i,weathered}} - 1, \tag{S18}$$

Here, C is the concentration of a base cation (j) or an immobile element (i) in the parent or weathered rock. When τ at the top of the weathering profile is 0, the composition of the weathering material is the same with respect to base cations and immobile element i as the parent and none of these elements have been lost to solution (they will be eroded instead of chemically weathered). When τ = -1, all of the element has been lost to solution and none is left to erode away.

Using the variables γ_{rock}, ζ_{rock}, and τ, we now define κ_{rock}, the long-term CO_2 sequestration coefficient of the rock:

$$\kappa_{rock} = \frac{1}{2}\tau_{silicate\ cations}(1-\gamma_{rock}) - \frac{1}{2}\zeta_{rock}. \tag{S19}$$

Here, (1- γ_{rock}) is the proportion of base cation equivalents associated with silicate minerals. We multiply this by 0.5 because 1 mol of CO_2 is sequestered during weathering of 2 eq of base cations when considered over 10^5 to 10^6 yr timescales (see Fig. 1 reactions 1 and 2). If pyrite oxidation is coupled to carbonate dissolution, 2 mols of CO_2 are released per mole of pyrite in the rock (see Fig. 1 reaction 7), yielding the term $-\frac{1}{2}\zeta_{rock}$ based on eq S16. Likewise, pyrite oxidation could be coupled to silicate dissolution. In this case, 1 mol of pyrite consumes 2 mols of silicate minerals. Because 1 mol of Ca-silicate mineral sequesters 1 mol of CO_2 over 10^5 to 10^6 yr timescales (see Fig. 1 reactions 1 and 2), CO_2 sequestration is reduced by 2 mols of CO_2 per mole pyrite in the rock. Again, based on eq S16, this is equivalent to $-\frac{1}{2}\zeta_{rock}$. Lastly, $\tau_{silicate\ cations}$ is the mass transfer coefficient for base cations in silicates at the land surface. It ranges from 0 (no base cations in silicate minerals have been removed by dissolution) to -1 (all the base cations in silicate minerals have been removed by dissolution).

Finally, noting that $\tau_{silicate\ cations}$ is generally not reported, we must instead calculate it from τ, the mass transfer coefficient for total base cations in the bulk rock:

$$\tau = \tau_{silicate\ cations}(1 - \gamma_{rock}) + \tau_{carbonate\ cations}\gamma_{rock}, \qquad \text{(S20)}$$

Again, we emphasize wet climates and low-carbonate terrain and implicitly assume that all carbonates are fully dissolved at the land surface (i.e., $\tau_{carbonate\ cations} = -1$) to solve for $\tau_{silicate\ cations}$:

$$\tau_{silicate\ cations} = \frac{\tau + \gamma_{rock}}{(1 - \gamma_{rock})}. \qquad \text{(S21)}$$

Now we substitute eq. S21 in eq. S19 and simplify to the final equation 3 from the main text:

$$\kappa_{rock} = \frac{1}{2}(\tau + \gamma_{rock} + \zeta_{rock}),$$

When $\kappa_{rock} < 0$, the rock has sequestered CO_2 from the atmosphere over the residence time of the soil and when $\kappa_{rock} > 0$ the rock released CO_2.

Mathematically, this equation is only valid as long at $\tau < -\gamma_{rock}$. The minimum value of κ_{rock} is -0.5, which is a pure silicate rock dissolved only by CO_2. The maximum value of κ_{rock} is 0.25, which is a pure carbonate rock weathered only by sulfuric acid. It is mathematically impossible for $\kappa_{rock} < -0.5$; however, it is mathematically possible to have $\kappa_{rock} > 0.25$. In these situations, there is more sulfuric acid in the system than can be buffered by both carbonate and silicate weathering.

2.4 Lag-time Calculation

Using rain chemistry data from the National Atmospheric Deposition Program (NADP; http://nadp.slh.wisc.edu/) site PA42, we calculated the annual flux of sulfate into Shale Hills from wet deposition. We used the flux data to calculate a trend in wet deposition over time and then used the regression to calculate when 39.5 mmol m^{-2} yr^{-1} was deposited (i.e., 31 years prior to today). Next we added dry deposition as an input (estimated as 30% wet deposition; Lynch and Corbett; 1989), fit a new regression to wet+dry deposition over time, and recalculated the lag time (i.e., 19 years; Fig. 4C). Although not explicitly calculated here, Hubbard Brook also shows a lag in deposition to export on similar timescales, which is consistent with the excess sulfate export observed in other studies (Likens et al., 2002).

2.5 Mineral-derived Solute Concentrations

The contributions of ankerite and calcite to the Ca^{2+} budget were calculated using the composition of the appropriate endmember (deep flowpath for SH and ER). Based on the stoichiometry of ankerite at Shale Hills and assuming all Mg^{2+} in deep flowpath water derives from ankerite, the concentration of Ca^{2+} from ankerite in any given stream sample, $[\mathrm{Ca}^{2+}]_{\mathrm{ankerite}}$, is calculated using the following equation.

$$[Ca^{2+}]_{ankerite} = 1.6\left(\frac{Mg^{2+}}{SO_4^{2-}}\right)_{deep} \alpha_{deep}[SO_4^{2-}]_{Total}, \tag{S22}$$

Here, 1.6 is the stoichiometric number relating Mg^{2+} to Ca^{2+} in Ankerite (see Table S3). $[Ca^{2+}]_{calcite}$ is calculated as the difference between the total Ca^{2+} and the ankerite-derived Ca^{2+}.

Similarly, the contributions of chlorite and illite to the Mg^{2+} budget were calculated using the composition of the appropriate endmember (shallow flowpath for SH). Based on the stoichiometry of illite at Shale Hills and assuming all K^+ in shallow flowpath water derives from illite, the concentration of Mg^{2+} from illite in any given sample, $[Mg^{2+}]_{illite}$, is calculated using the following equation:

$$[Mg^{2+}]_{illite} = 0.28\left(\frac{K^{+}}{SO_4^{2-}}\right)_{shallow} \alpha_{shallow}[SO_4^{2-}]_{Total}, \tag{S23}$$

Here, 0.28 is the stoichiometric number relating K^+ to Mg^{2+} in illite (see Table S3). The concentration of chlorite-derived Mg^{2+} is calculated as the difference between the total Mg^{2+}, the ankerite-derived Mg^{2+} and the illite-derived Mg^{2+}. Fluxes of solutes derived from each mineral are summarized in Table S5 for Shale Hills.

Section 3 Seasonality of Pyrite-sulfate Fluxes

At Shale Hills, the proportion of pyrite-derived sulfate leaving the catchment accounts for 23% of the annual sulfate flux (Table 1) but ranges from 99% of total sulfate in the dry season (summer, fall) to as low as 3% in the wet season (winter, spring, Fig. 5A). This is easily explained because the stream is sustained by deep groundwater that flows up into the stream from the deep pyrite reaction front during the dry summer and fall but not in the winter and less acid rain enters the catchment in the dry season (Li et al., 2017).

Section 4 Rain-correction

For simplicity, we do not correct $[\Sigma^+]_{total}$ (eq. S3) for rain chemistry; however, it is likely that some of the cations in the stream are derived from rain, rather than weathering. Because Cl in the stream is only derived from precipitation, we can apply a basic correction by subtracting cations from $[\Sigma^+]_{total}$ that balance the Cl in the stream water. Because all of our calculations rely on $[\Sigma^+]_{total}$, and individual cations, it does not matter which cations are subtracted to balance the Cl. The correction results in very little change in the calculated fluxes and does not change the interpretations of the study (see Table S6).

Table S1. Endmember compositions for synthetic dataset

	Endmember1	Endmember2
$[Ca^{2+}]/[SO_4^{2-}]$[a]	8±0.7	0±0.2
$[Mg^{2+}]/[SO_4^{2-}]$	3±0.3	0.5±0.1
$[Na^{+}]/[SO_4^{2-}]$	2±0.1	4±0.2
$[K^{+}]/[SO_4^{2-}]$	2±0.3	1±0.1
$[Cl^{-}]/[SO_4^{2-}]$	0±0.1	5±0.6

a) All analytes are reported in molar concentration ratios.

Table S2. Chemical composition and interpretations of end members from NMF model (see also Figure S3)

	Shale Hills		**East River**		**Hubbard Brook**		
Component	1	2	1	2	1	2	3
$[\mathrm{Ca}^{2+}]/[\mathrm{SO}_4^{2-}]$[a]	0.0±0.0	10.0±3.6	1.6±0.6	3.2±0.6	0.2±0.1	1.0±0.2	0.2±0.1
$[\mathrm{Mg}^{2+}]/[\mathrm{SO}_4^{2-}]$	1.6±0.7	2.9±1.0	0.4±0.1	0.7±0.1	0.2±0.1	0.4±0.1	0.3±0.1
$[\mathrm{Na}^{+}]/[\mathrm{SO}_4^{2-}]$	0.5±0.2	0.6±0.2	0.1±0.0	0.2±0.0	2.1±0.5	0.4±0.2	0.4±0.2
$[\mathrm{K}^{+}]/[\mathrm{SO}_4^{2-}]$	0.6±0.3	0.7±0.2	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.5±0.1
$[\mathrm{Cl}^{-}]/[\mathrm{SO}_4^{2-}]$	1.6±0.7	0.0±0.0	0.1±0.0	0.0±0.0	0.4±0.1	0.2±0.1	0.4±0.1
Interpretation	Shallow	Deep	Shallow	Deep	Shallow Till	Moderately Shallow Till	Deep

a) All analytes are reported in molar concentration ratios.

Table S3. Mineral reactions with CO_2 and H_2SO_4

Reaction	Equation
1	$\mathrm{Calcite} + CO_2 + H_2O \rightarrow Ca^{2+} + 2HCO_3^-$
2	$\mathrm{Dolomite} + 2CO_2 + 2H_2O \rightarrow Ca^{2+} + Mg^{2+} + 4HCO_3^-$
3	$\mathrm{Ankerite} + 2CO_2 + 2H_2O \rightarrow Ca^{2+} + 0.62Mg^{2+} + 4HCO_3^-$
4	$2\mathrm{Calcite} + H_2SO_4 \rightarrow 2Ca^{2+} + 2HCO_3^- + SO_4^{2-}$
5	$\mathrm{Dolomite} + H_2SO_4 \rightarrow Ca^{2+} + Mg^{2+} + 2HCO_3^- + SO_4^{2-}$
6	$\mathrm{Ankerite} + H_2SO_4 \rightarrow Ca^{2+} + 0.62Mg^{2+} + 2HCO_3^- + SO_4^{2-}$
7[a]	$\mathrm{Chlorite} + 0.6O_2 + 1.2CO_2 + 1.2H_2O \rightarrow 1.2\mathrm{Hematite} + \mathrm{Vermiculite} + 0.6Mg^{2+} + 3.6H_2O$
8[a]	$\mathrm{Illite} + 0.91CO_2 + 4.15H_2O \rightarrow 1.08\mathrm{Kaolinite} + 0.48\ \mathrm{Goethite} + 0.07Mg^{2+} + 0.77K^+ + 1.15H_4SiO_4 + 0.91\ HCO_3^-$

Calcite: $CaCO_3$
Dolomite: $CaMg(CO_3)_2$
Ankerite: $Ca(Fe_{0.34}Mg_{0.62}Mn_{0.04})(CO_3)_2$
Chlorite: $(Fe^{2+}_{0.40}Mg_{0.15}Al_{0.35})_6(Si0_{.76}Al_{0.24})_4O_{10}(OH)_8$
Illite: $K_{0.69}(Si_{3.24}Al_{0.76})(Al_{1.69}Fe^{3+}_{0.10}Fe^{2+}_{0.16}Mg_{0.19})O_{10}(OH)_2$
Hematite: Fe_2O_3
Vermiculite: $(Mg_{0.3}Al_{2.1})(Si_{0.76}Al_{0.24})_4O_{10}(OH)_2$
Kaolinite: $Al_2Si_3O_{10}(OH)_2$
Goethite: FeOOH

a) Mineral stoichiometries for chlorite and illite are reported in Sullivan et al. (2016)

Table S4. Relevant element concentrations and parameters to determine κ_{rock}

	Shale Hills[a]		East River[b]		Hubbard Brook[c]	
(meq/kg)	Mean	sd	Mean	sd	Mean	sd[h]
Cation concentrations in parent	2309	439	4003	999	3321	664
Cation concentrations in topsoil	1368	552	1810	263	1528	306
Total sulfur in parent rock	100	19	686	312	119	24
Inorganic carbon in parent rock	250	42	1083	417	42	8
γ_{rock}	0.22	0.05	0.54	0.25	0.03	0.01
ζ_{rock}	0.04	0.01	0.17	0.09	0.04	0.01
τ^{g}	-0.43[d]	0.19	-0.55[e]	0.16	-0.45[f]	0.13
$\tau_{silicate\ cations}$	-0.27	0.26	-0.01	0.64	-0.43	0.18
κ_{rock}	-0.08	0.11	0.08	0.17	-0.19	0.11

[a] Values from Gu et al. (2020b)

[b] Values from Wan et al. (2019)

[c] Values from Johnson et al. (1968) and Bailey et al. (2004)

[d] τ calculated as averages from samples taken at the land surface for bulk composition data for 3 boreholes

[e] τ calculated from the average of the top 8 cm of 5 cores as reported in Wan et al. (2019)

[f] τ calculated from unweathered schist and reported soil data in Johnson et al. (1968)

[g] immobile element used in τ calculations is Ti

[h] No error on measurements were reported (Johnson et al., 1968); we therefore assumed 20% error.

Table S5. Fluxes of SO_4^{2-}, Ca^{2+}, and Mg^{2+} by contributor at Shale Hills

Analyte	Fraction	Flux (mmol m^{-2} yr^{-1})
SO_4^{2-}	Total	50.3 ± 0.3
	Rain-derived	38.9 ± 1.0 (77%)[a]
	Pyrite	11.2 ± 0.9 (23%)
Ca^{2+}	Total	99.5 ± 15.2
	Calcite	56.1 ± 8.4 (56%)
	Ankerite	43.3 ± 6.9 (44%)
Mg^{2+}	Total	51.8 ± 7.4
	Ankerite	28.2 ± 4.4 (54%)
	Chlorite	21.3 ± 5.7 (41%)
	Illite	2.4 ± 0.6 (5%)

[a] Number in parentheses is the percent of the total flux for that element

Table S6. Fluxes of SO_4^{2-}, Cations, and CO_2 and CO_2 Sequestration Coefficients Corrected for Rain Cl

	Shale Hills		**East River**		**Hubbard Brook**	
	Base Cation Fluxes (meq m^{-2} yr^{-1})					
	Uncorrected	*Corrected*	*Uncorrected*	*Corrected*	*Uncorrected*	*Corrected*
Total base cation flux	336 ± 13	316 ± 13	1540 ± 30	1530 ± 30	84.6 ± 0.8	74.4 ± 0.8
Base cation flux from CO_2-weathering of silicates	12.6 ± 21.1	5.0 ± 21.1	315 ± 58	300 ± 58	24.1± 0.8	14.8 ± 0.8
Base cation flux from CO_2-weathering of carbonates	216 ± 16	216 ± 16	587 ± 48	587 ± 48	-	-
Base cation flux from H_2SO_4-weathering of silicates	62.4 ± 1.0	66.0 ± 1.0	152 ± 4	152 ± 4	60.5 ± 0.2	59.5 ± 0.2
Base cation flux from H_2SO_4-weathering of carbonates	44.8 ± 1.9	44.8 ± 1.9	488 ± 9	488 ± 9	-	-
	Fluxes (mmol m^{-2} yr^{-1})					
	Uncorrected	*Corrected*	*Uncorrected*	*Corrected*	*Uncorrected*	*Corrected*
Total sulfate flux	50.3 ± 0.3		197.5 ± 1.0		30.3 ± 0.1	
Sulfide-derived sulfate flux	11.2 ± 0.9		122.1 ± 4.3		9.1 ± 0.1	
Rain-derived sulfate flux	38.9 ± 1.0		75.9 ± 4.2		21.2 ± 0.6	
CO_2 sequestration or release	4.9 ± 10.7	8.7 ± 10.7	-35.6 ± 30.4	-27.9 ± 30.4	-12.1 ± 0.4	-7.4 ± 0.4
	CO_2 Sequestration Coefficients					
	Uncorrected	*Corrected*	*Uncorrected*	*Corrected*	*Uncorrected*	*Corrected*
κ_{stream}	0.01 ± 0.03	0.03 ± 0.03	-0.02 ± 0.02	-0.02 ± 0.02	-0.14 ± 0.01	-0.10 ± 0.01
κ_{rock}	-0.08 ± 0.11		0.08 ± 0.17		-0.19 ± 0.02	

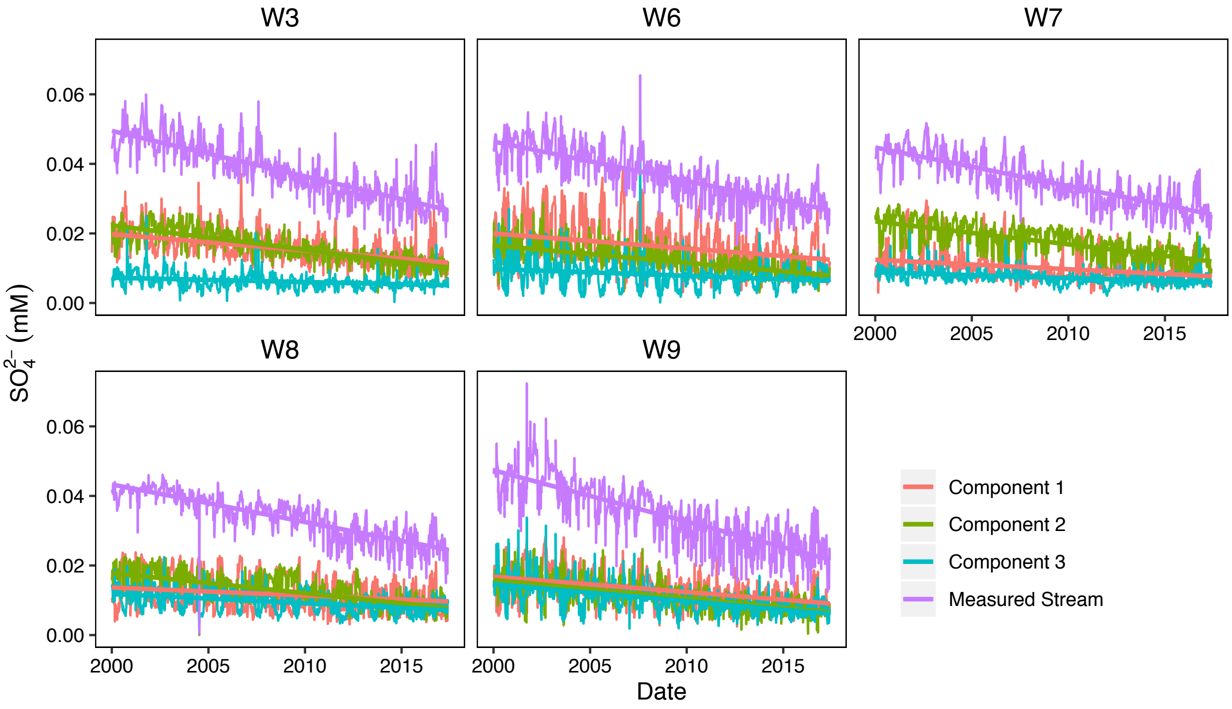

Figure S1. Time series showing [SO_4^{2-}]in stream water for the three components calculated from NMF for the 5 analyzed sub-catchments for Hubbard Brook. The measured stream sulfate concentrations (i.e. total sulfate) are also shown. Components 1 through 3 have been inferred to indicate weathering along flowpaths that are shallow, moderately shallow, and deep, respectively (see text and Table S2).

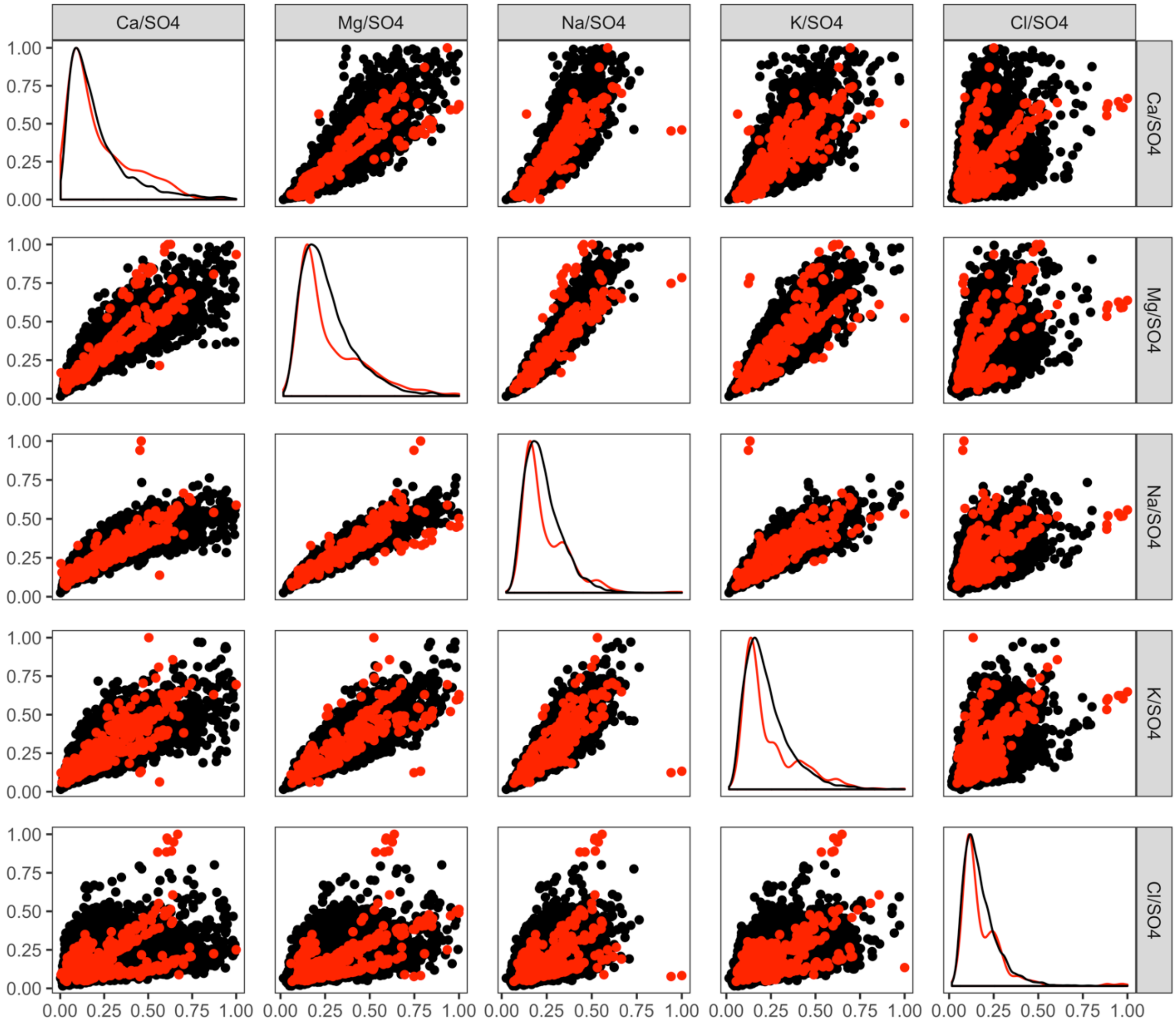

Figure S2. Matrix of plots showing measured concentration ratios in stream water at Shale Hills normalized to their maximum value (red) and the bootstrapped normalized concentration ratios (black). Off-diagonal plots show every combination of element ratio pairs to illustrate covariation in the dataset. Plots on the diagonal are element ratio distributions to illustrate that the bootstrapped dataset matches the distribution of the measured stream samples.

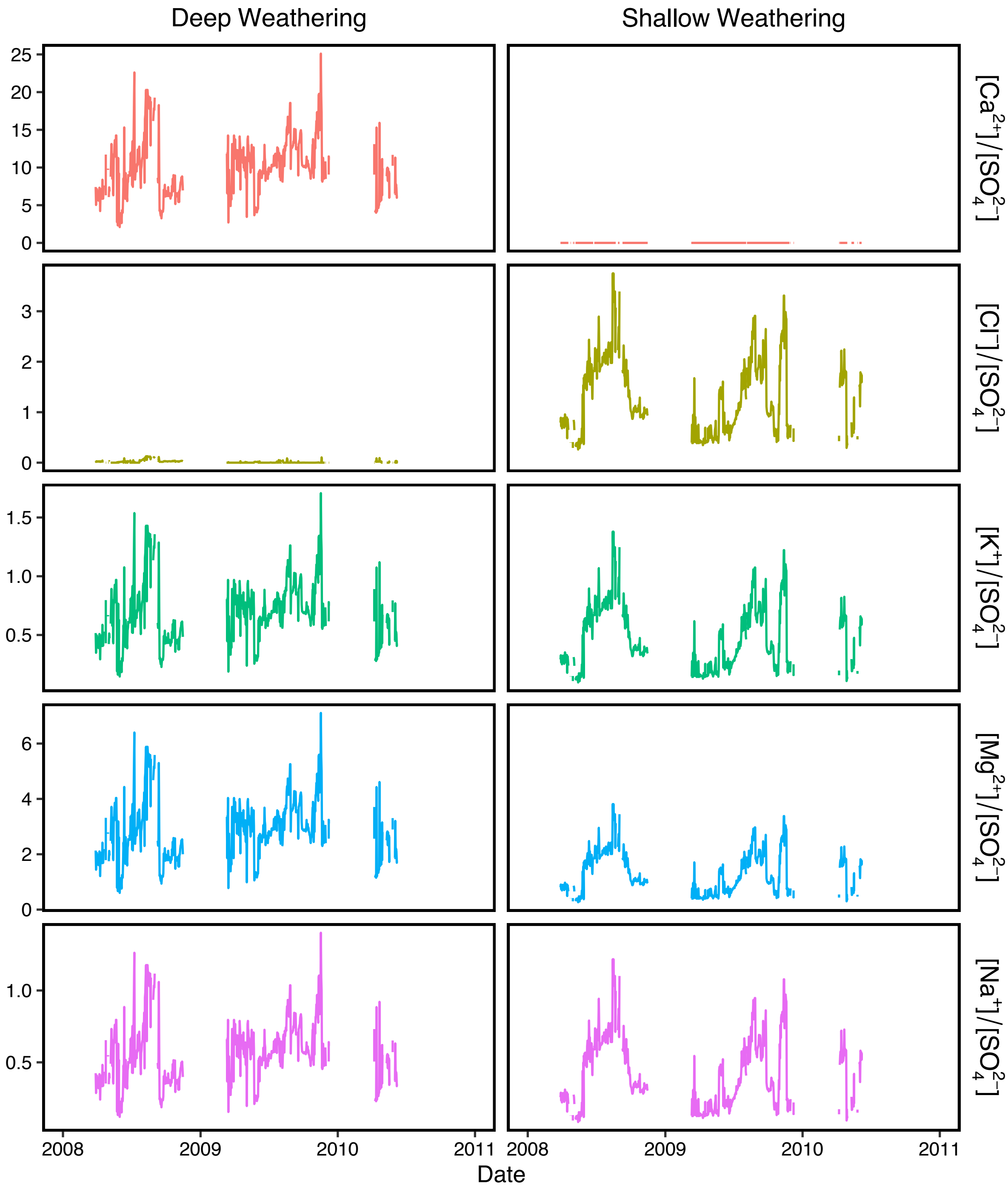

Figure S3. Plot showing the variation in end member composition over time for shallow and deep weathering end members at Shale Hills.

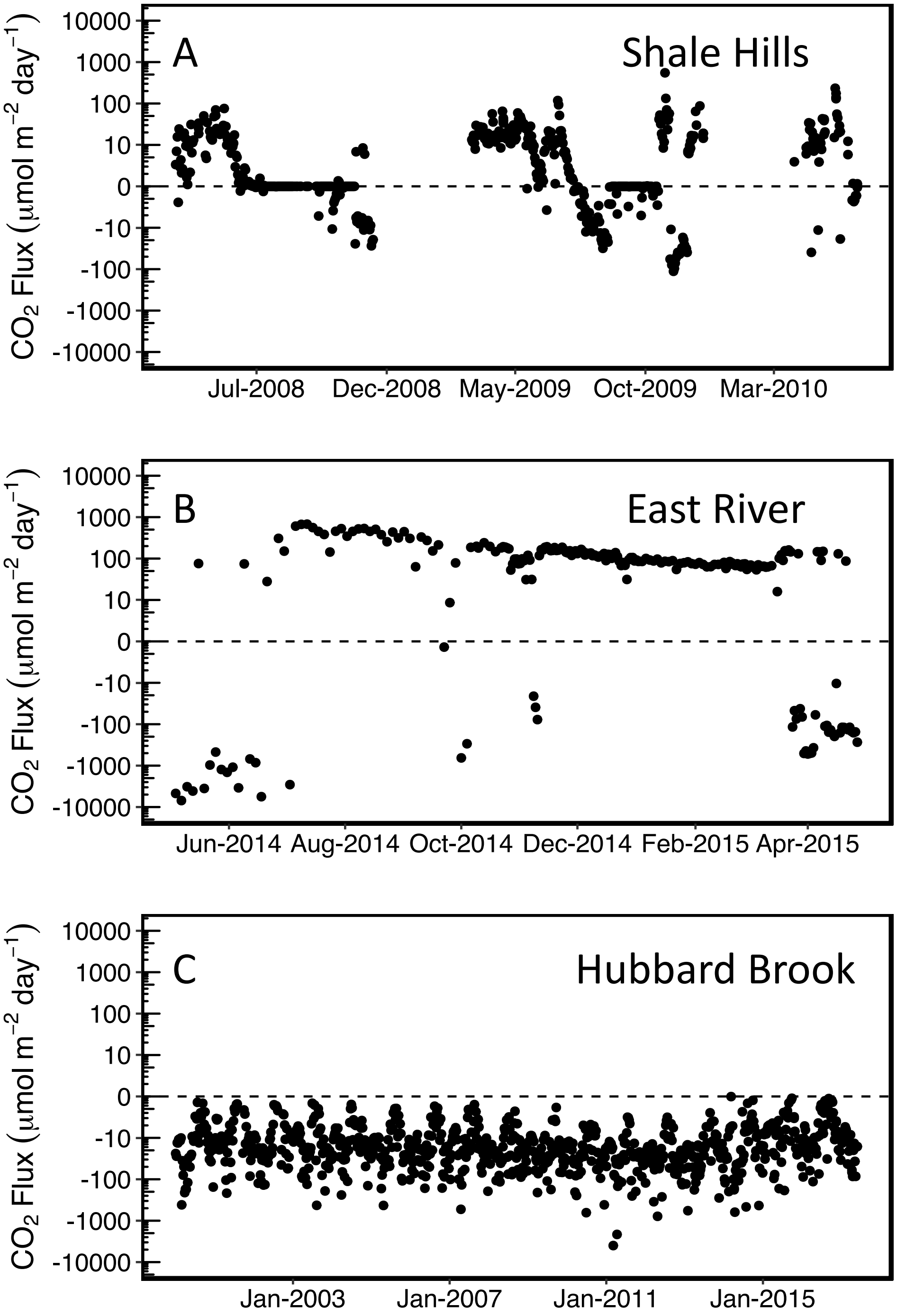

Figure S4. Plots showing the CO_2 fluxes at Shale Hills (A), East River (B), and Hubbard Brook (C) through weathering reactions inferred from the stream chemistry (see SM section 2.2).

References:

Johnson, N. M., Likens, G. E., Bormann, F. H., & Pierce, R. S. (1968). Rate of chemical weathering of silicate minerals in New Hampshire. *Geochimica et Cosmochimica Acta*, *32*(5), 531-545.

Lautz, L. K., et al. (2014). Using discriminant analysis to determine sources of salinity in shallow groundwater prior to hydraulic fracturing. *Environmental science & technology*, *48*(16), 9061-9069.

Lynch, J. A., & Corbett, E. S. (1989). Hydrologic control of sulfate mobility in a forested watershed. *Water Resources Research*, *25*(7), 1695-1703.

Moatar, F., Meybeck, M., Raymond, S., Birgand, F., & Curie, F. (2013). River flux uncertainties predicted by hydrological variability and riverine material behaviour. *Hydrological processes*, *27*(25), 3535-3546.

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, *12*(Oct), 2825-2830.

Riebe, C.S., Kirchner, J.W., Finkel, R.C., 2003. Long-term rates of chemical weathering and physical erosion from cosmogenic nuclides and geochemical mass balance. Geochimica et Cosmochimica Acta 67, 4411-4427.