Supplemental Material:

Machine Learning Deciphers CO₂ Sequestration and Subsurface Flowpaths from Stream Chemistry

Andrew R. Shaughnessy[1], Xin Gu[1], Tao Wen[2], Susan L. Brantley[1,2]

1. Department of Geosciences, Pennsylvania State University, University Park, PA, USA
2. Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA, USA.

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 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):

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

(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+}) at timepoint n, Q_n is the discharge measured at timepoint n, $\bar{q}$ is the mean annual discharge (i.e., sum of daily discharge measurements/number of daily discharge measurements), A is the basin area (Shale Hills =

45   8.0x10^4 m^2; East River = 8.5x10^7 m^2; Hubbard Brook = 4.0x10^5 m^2 (W3), 1.5x10^5 m^2 (W6), 7.7x10^5 m^2 (W7), 6.1x10^5 m^2 (W8), 7.0x10^5 m^2 (W9)), and β is a coefficient for unit conversions to mmol m^{-2} yr^{-1} (i.e., 3.15 x10^{10} l s m^{-3} yr^{-1}) or to meq m^{-2} yr^{-1} (i.e., 3.15 x10^{10} l s m^{-3} yr^{-1} multiplied by species charge (meq/mmol)).

S2.2 Using Stream Chemistry to Calculate CO$_2$ Drawdown or Release

50   To calculate the fluxes of CO$_2$ drawdown or release from weathering, we identify 4 weathering reactions: 1) CO$_2$-driven weathering (CO$_2$-weathering) of silicates, 2) H$_2$SO$_4$-driven weathering (H$_2$SO$_4$-weathering) of silicates, 3) CO$_2$-weathering of carbonates, and 4) H$_2$SO$_4$-weathering of carbonates. From the NMF decomposition described above and in the main body of the manuscript, all of these reactions are constrained.

For example, NMF yields proportions of sulfate derived from each endmember (deep flow, shallow flow,

55   and sometimes medium-depth flow) for each sample. We can easily constrain the total concentration of base cations derived from carbonate minerals as the Ca and Mg concentrations in the deep weathering endmember for either of the two shale watersheds. The total concentration of base cations minus the base cations derived from carbonate minerals is the concentration of silicate-derived base cations. Field observations tell us that that carbonate and pyrite reaction fronts are co-located. We constrain H$_2$SO$_4$-weathering of carbonates (i.e., coupling of pyrite oxidation to

60   carbonate dissolution) using the concentration of sulfate in the deep weathering endmember from the NMF decomposition. We assume that all of the sulfate in the deep weathering endmember dissolves carbonate minerals. Any carbonate-derived base cations that are in excess of what could have been produced by the sulfuric acid in the deep weathering endmember are associated with CO$_2$-weathering of carbonates. Similarly, because there are no carbonate minerals in the shallow subsurface, we assume that all of the sulfuric acid associated with acid rain

65   dissolves silicate minerals. Any cations not associated with H$_2$SO$_4$-weathering of silicates we assume to be associated with CO$_2$-weathering of silicates. For Hubbard Brook, because there are no carbonate minerals, all of the sulfate is associated with the dissolution of silicate minerals.

For long-term CO$_2$ dynamics, we consider H$_2$SO$_4$-weathering of silicates and CO$_2$-weathering of carbonates to be CO$_2$ neutral, while CO$_2$-weathering of silicates sequesters CO$_2$ and H$_2$SO$_4$-weathering of carbonates releases

70   CO$_2$ (Fig. 1). We calculate the potential sequestration or release of CO$_2$ in a sample as 0.5 times the base cation concentration from CO$_2$-weathering of silicates (on a charge equivalent basis) minus the product of 0.25 times the base cation concentration from H$_2$SO$_4$-weathering of carbonates (on a charge equivalent basis) (Fig. 1). Next, we calculate the flux of CO$_2$ following eq. S2 in section S2.1.

S2.3 Rock CO$_2$ Calculation

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. The three most important factors are 1) the ratio of base cations in carbonates relative to silicates, 2) the ratio of acid-generating units of pyrite relative to total base cations in carbonates+ silicate minerals, and 3) 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, 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). We note that for this calculation we do not include organic carbon. For (1), we define the carbonate/silicate factor, γ_{rock}, which is the proportion of base cations in the rock derived from carbonate minerals divided by the total base cations:

$$\gamma_{rock} = \frac{2C_{Ca,carb}+2C_{Mg,carb}}{2C_{Ca,Total}+2C_{Mg,Total}+C_{Na,Total}+C_{K,Total}}, \tag{S3}$$

Here $C_{X,k}$ is the molar concentration of base cation (X = Ca, Mg, Na, or K) in carbonates (k = carb) or in carbonate + silicate minerals (k = Total). Eqn. (S3) expresses γ_{rock} as proton equivalents/kg rock. Often, rather than using quantitative X-Ray Diffraction (XRD), inorganic carbon (IC) in the rock is measured via acid titration to determine the amount of carbonate in the rock. Therefore, equation S3 can be re-written in terms of IC as:

$$\gamma_{rock} = \frac{2C_{IC}}{2C_{Ca,Total}+2C_{Mg,Total}+C_{Na,Total}+C_{K,Total}}, \tag{S4}$$

Here, C_{IC} refers to concentration of inorganic carbon. By definition, γ_{rock} ranges from 0 (where all cations derive from silicates) to 1 (where all cations derive from carbonates). Likewise, 1- γ_{rock} is the proportion of cations derived from silicate minerals.

When pyrite oxidizes it produces sulfuric acid that can dissolve carbonate and silicate minerals. This impacts CO$_2$ dynamics by either releasing CO$_2$ (H$_2$SO$_4$-weathering of carbonates) or by diminishing the silicate content of the rock (H$_2$SO$_4$-weathering of silicates), thereby diminishing the rock's capacity to sequester CO$_2$. Here, we define a new variable, ζ_{rock}, which is the acid generation capacity relative to the base cations in the rock expressed on a proton equivalents basis:

$$\zeta_{rock} = \frac{4C_{py}}{2C_{Ca,Total}+2C_{Mg,Total}+C_{Na,Total}+C_{K,Total}}, \tag{S5}$$

Here, the subscript *py* refers to pyrite (proton equivalents/kg rock). We multiply the concentration of pyrite (i.e., C_{py}; mol/kg) by 4 (eq/mol) because 4 moles of acid are produced (as H$_2$SO$_4$) per mole of pyrite as shown in reaction S6.

$$FeS_2 + \frac{15}{4}O_2 + 2H_2O \rightarrow \frac{1}{2}Fe_2O_3 + 2H_2SO_4, \tag{S6}$$

Lastly, in many catchments, the bulk chemistry of parent rock is not indicative of the CO$_2$ sequestration because silicate minerals are kinetically slow to dissolve and they do not completely dissolve before the rock physically erodes. (Here we assume all carbonate minerals chemically weather away before exposure at land surface,

i.e. wet climates and non-karst terrains). 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.

$$\tau_{i,j} = \frac{C_{j,weathered}C_{i,parent}}{C_{j,parent}C_{i,weathered}} - 1, \tag{S7}$$

110   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 the element has been lost to solution before being eroded away. When τ is -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

115   rock.

$$\kappa_{rock} = \frac{1}{2}\tau_{silicate\ cations}(1 - \gamma_{rock}) - \frac{1}{2}\zeta_{rock}, \tag{S8}$$

Here, (1- γ_{rock}) is the proportion of base cations associated with silicate minerals. We multiply this by 0.5 because 1 eq of CO_2 is sequestered per 2 eq of base cations weathered. If pyrite oxidation is coupled to carbonate dissolution, 2 eq of CO_2 are released per mole of pyrite in the rock. Based on our definition of ζ_{rock}, this is equivalent to $\frac{1}{2}\zeta_{rock}$. If

120   pyrite oxidation is coupled to silicate dissolution, then CO_2 sequestration is reduced by 2 eq of CO_2 per mole pyrite in the rock. Again, based on eq 8, this is equivalent to $\frac{1}{2}\zeta_{rock}$. Lastly, $\tau_{silicate\ cations}$ is the mass transfer coefficient for silicate cations at the land surface. It ranges from 0 (no silicate minerals dissolve) to -1 (all the silicate minerals dissolve).

        Finally, noting that $\tau_{silicate\ cations}$ is generally not reported just for silicate mineral base cations, we must instead

125   calculate it from τ, the mass transfer coefficient for total base cations in the bulk rock:

$$\tau = \tau_{silicate\ cations}(1 - \gamma_{rock}) + \tau_{carbonate\ cations}\gamma_{rock}, \tag{S9}$$

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

130     $$\tau_{silicate\ cations} = \frac{\tau + \gamma_{rock}}{(1 - \gamma_{rock})}, \tag{S10}$$

Now we substitute eq. S10 in eq. S8 and simplify to the final equation 2 from the main text:

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

When $\kappa_{rock} < 0$, the rock has sequestered CO_2 over the weathering duration and when $\kappa_{rock} > 0$ the rock on net

135   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. 3C). 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 this endmember derives from ankerite, the concentration of Ca^{2+} from ankerite in any given sample, $[Ca^{2+}]_{ankerite}$, is calculated using the following equation.

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

(S11)

Here, Mg^{2+}/SO$_4^{2-}$ is the model-derived ratio of Mg to SO4 from the deep weathering endmember for that sample (subscript deep), α_{deep} is the sulfate mixing proportion for the deep weathering endmember for that sample, [SO$_4^{2-}$]$_{Total}$ is the total concentration of sulfate in that sample, and 1.6 is the stoichiometric number relating Mg^{2+} to Ca^{2+} in Ankerite (see Table S2). The concentration of calcite-derived Ca^{2+} 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 this endmember derives from illite, the concentration of Mg^{2+} from illite in any given sample, $[Mg^{2+}]_{illite}$, is calculated using the following equation.

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

(S12)

Here, K$^+$/SO$_4^{2-}$ is the model-derived ratio of K to SO4 from the shallow weathering endmember for that sample (subscript *shallow*), $\alpha_{shallow}$ is the sulfate mixing proportion for the shallow weathering endmember for that sample, [SO$_4^{2-}$]$_{Total}$ is the total concentration of sulfate in that sample, and 0.28 is the stoichiometric number relating K$^+$ to Mg^{2+} in illite (see Table S2). 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.

175

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 up to 99% of total sulfate in the dry season (summer, fall) and to as low as 3% in the wet season (winter, spring, Fig. 3A). This is easily explained because the stream is sustained by deep groundwater that

180 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).

185

Table S1. Chemical composition and geochemical interpretations of the end members derived from the NMF model

Site	Shale Hills		East River		Hubbard Brook		
Component	1	2	1	2	1	2	3
Ca/SO4	0.0[a]	10.0±3.6	1.6±0.6	3.2±0.6	0.2±0.1	1.0±0.2	0.2±0.1
Mg/SO4	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
Na/SO4	0.5±0.2	0.6±0.2	0.1	0.2	2.1±0.5	0.4±0.2	0.4±0.2
K/SO4	0.6±0.3	0.7±0.2	0.0	0.0	0.0	0.0	0.5±0.1
Cl/SO4	1.6±0.7	0.0	0.1	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) Numbers without errors have negligible variability in the model output

190 **Table S2. Mineral reactions with CO_2 and H_2SO_4**

Reaction	Equation
1	Calcite + CO_2 + H_2O → Ca^{2+} + $2HCO_3^-$
2	Dolomite + $2CO_2$ + $2H_2O$ → Ca^{2+} + Mg^{2+} + $4HCO_3^-$
3	Ankerite + $2CO_2$ + $2H_2O$ → Ca^{2+} + $0.62Mg^{2+}$ + $4HCO_3^-$
4	2Calcite + H_2SO_4 → $2Ca^{2+}$ + $2HCO_3^-$ + SO_4^{2-}
5	Dolomite + H_2SO_4 → Ca^{2+} + Mg^{2+} + $2HCO_3^-$ + SO_4^{2-}
6	Ankerite + H_2SO_4 → Ca^{2+} + $0.62Mg^{2+}$ + $2HCO_3^-$ + SO_4^{2-}
7	Chlorite + $0.6O_2$ + $1.2CO_2$ + $1.2H_2O$ → 1.2Hematite + Vermiculite + $0.6Mg^{2+}$ + $3.6H_2O$
8	Illite + $0.91CO_2$ + $4.15H_2O$ → 1.08Kaolinite + 0.48 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: $(Fe^{2+}_{0.40}Mg_{0.15}Al_{0.35})_6(Si0_{.76}Al_{0.24})_4O_{10}(OH)_8$
195 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

200

Table S3. Relevant element concentrations and parameters to determine CO$_2$ dynamics

(meq/kg)	Shale Hills[a]		East River[b]		Hubbard Brook[c]	
	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., *Accepted*
[b]Values from Wan et al. (2019)
[c]Values from Johnson et al. (1968) and Bailey et al. (2004).
[d]τ calculated from samples at the land surface 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 reported (Johnson et al., 1968), we assume 20% error.

Table S4. Endmember compositions for synthetic dataset

	Endmember1	Endmember2
Ca/SO4	8±0.7	0±0.2
Mg/SO4	3±0.3	0.5±0.1
Na/SO4	2±0.1	4±0.2
K/SO4	2±0.3	1±0.1
Cl/SO4	0±0.1	5±0.6

215 **Table S5. Fluxes of SO_4^{2-}, Ca^{2+}, and Mg^{2+} by contributor**

Site	Analyte	Fraction	Flux (mmol m^{-2} yr^{-1})
Shale Hills	SO_4^{2-}	Total	51.8 ± 12.1
		Rain-derived	39.5 ± 10.8 (77%)[a]
		Pyrite	12.1 ± 3.1 (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%)
East River	SO_4^{2-}	Total	305.2
		Pyrite	219.2 ± 33.8 (62%)
Hubbard Brook	SO_4^{2-}	Total	29.5 ± 2.1
		Pyrrhotite	8.6 ± 1.8 (29%)

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

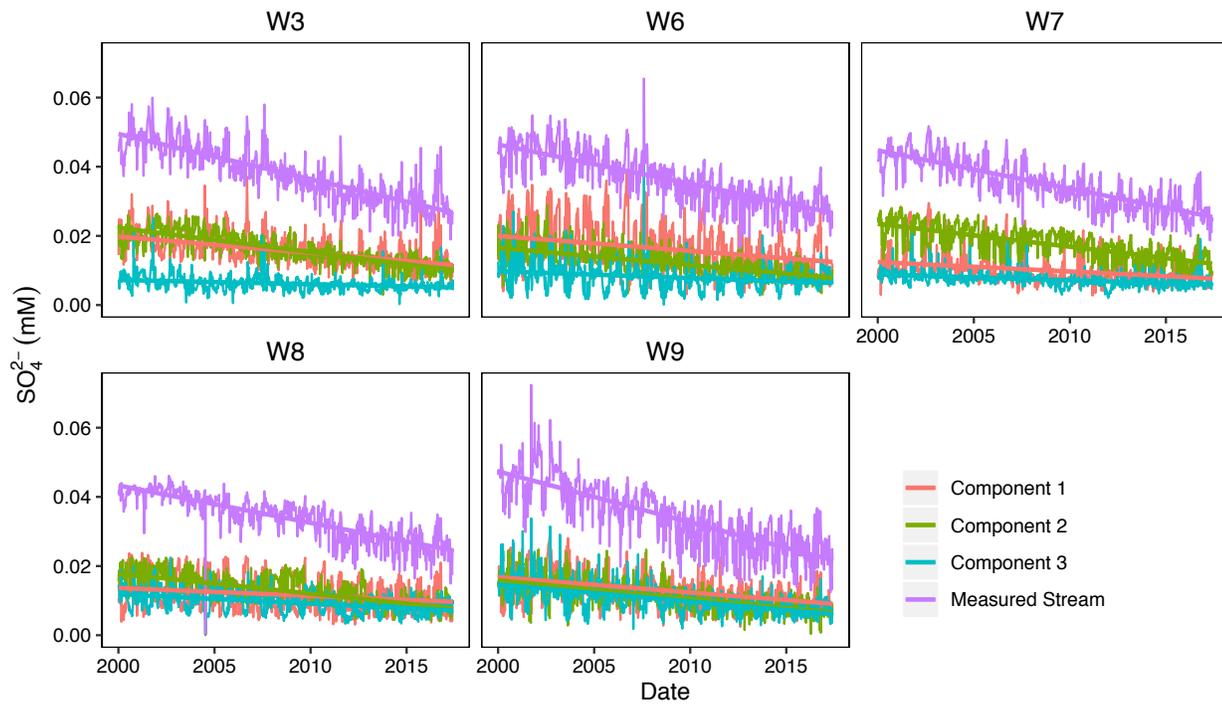



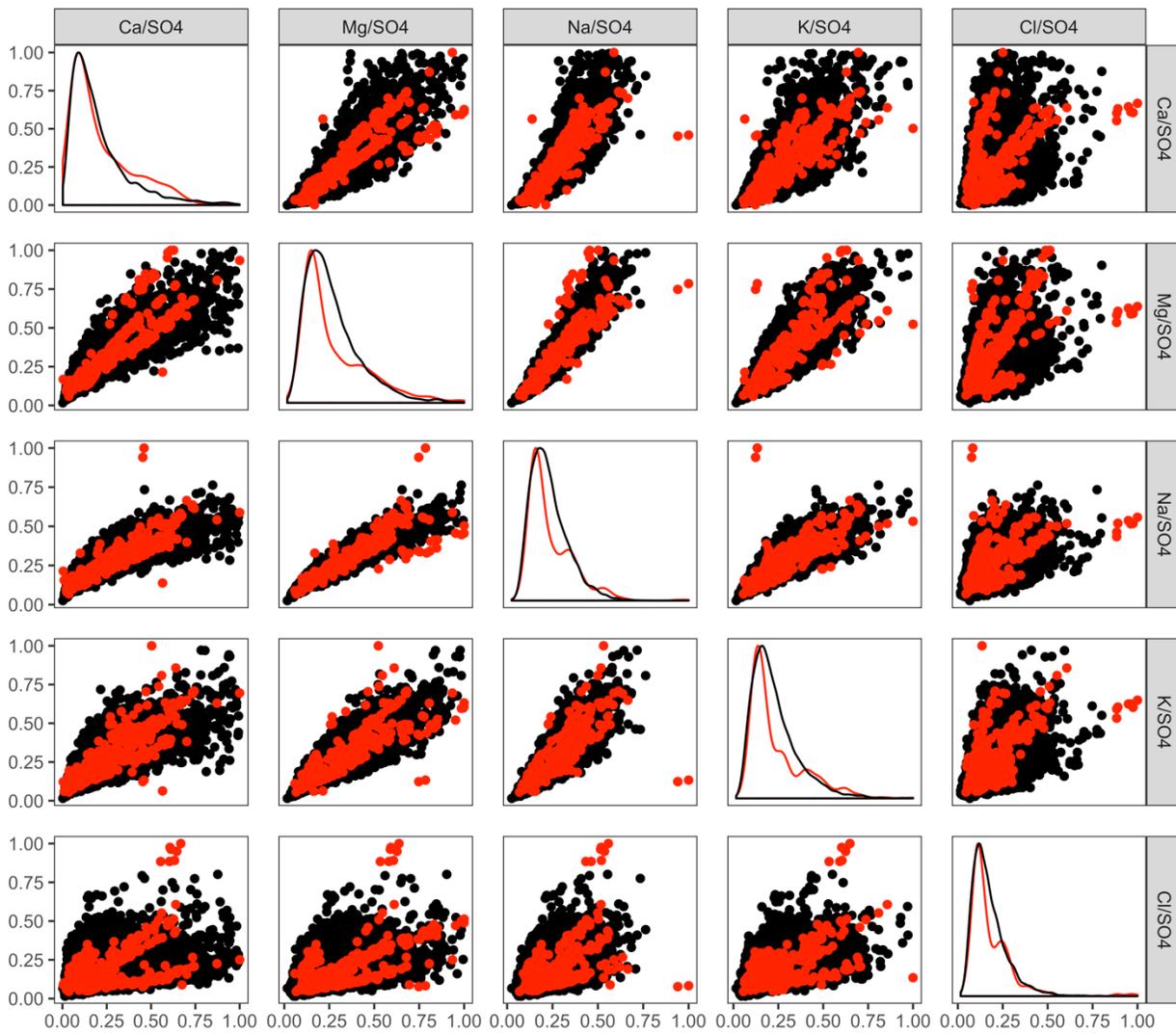
Figure S1. Time series showing the concentration of sulfate in stream water for the three components calculated from NMF for the 5 analyzed sub-catchments for Hubbard Brook. The measured stream sulfate concentrations (i.e. total sulfate) are also shown. Components 1 through 3 have been inferred to indicate weathering of soils on till, moderately shallow till weathering, and bedrock weathering, respectively (see text and Table S1).

225

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.

230

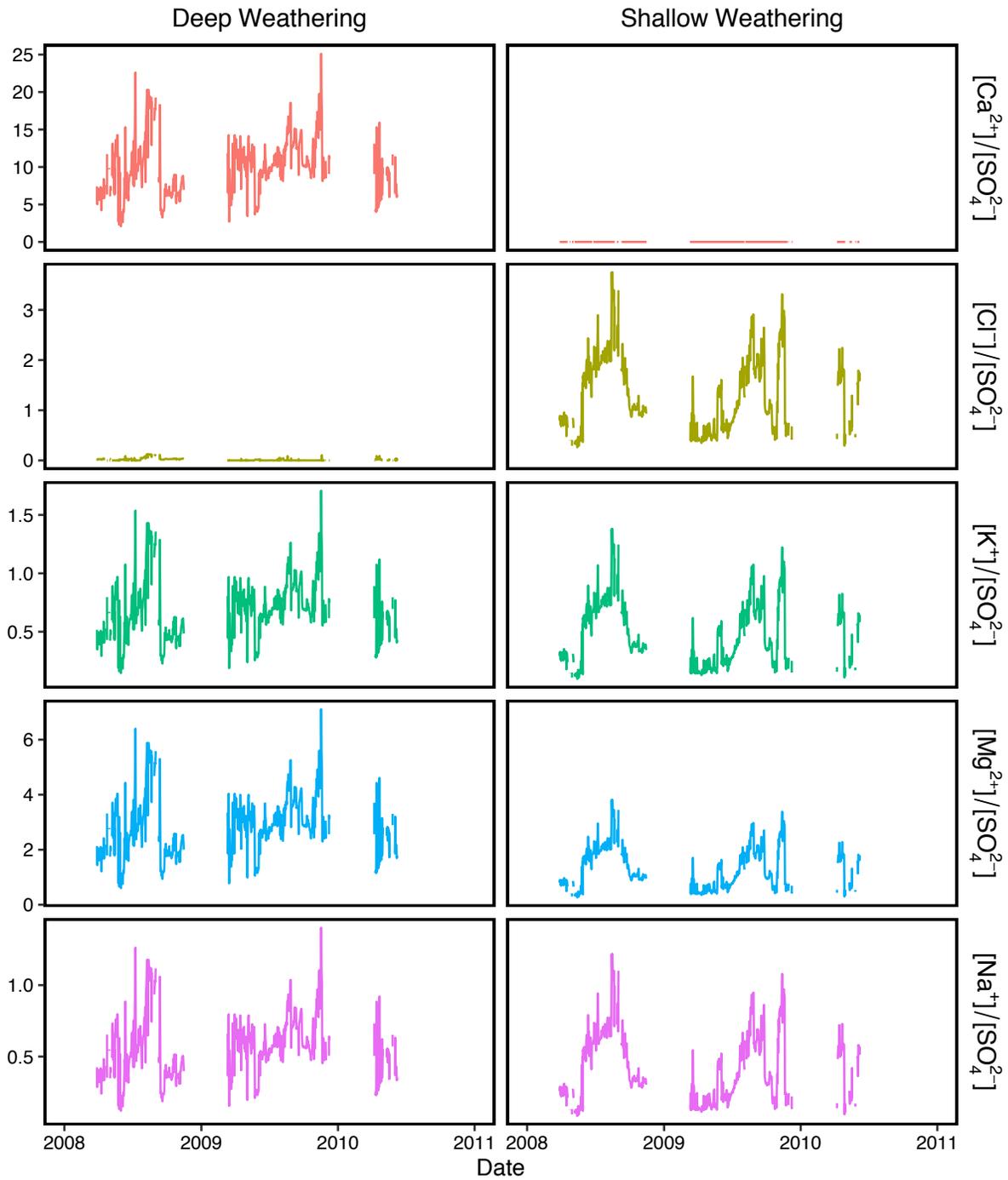



Figure S3. Plot showing the variation in end member composition over time for shallow and deep weathering end members at Shale Hills.

235

References:

Gu. X. et al. (2020). Seismic refraction tracks porosity generation and possible CO2 production at depth under a headwater catchment. *PNAS. Accepted.*

240   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.

245

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

250   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.

255   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.