



Implementation of a variance-targeting sum-constrained Gaussian sampling algorithm in the MEANDIR (v1.4) river inversion model

Preston Cosslett Kemeny¹, Ofer Grossman²

¹ Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

² Grossman Mathematical Consulting, Ithaca, NY, USA

Correspondence to: Preston Cosslett Kemeny (preston.kemeny@whoi.edu)

Abstract. Quantitative analysis of dissolved river chemistry underlies study of the chemical weathering reactions that stabilize the global carbon cycle over geological timescales. Because different chemical sources have distinct implications for the impact of weathering on biogeochemical cycles, interpreting fluvial chemistry requires constraining the fraction of each solute derived from separate lithologic or process end-members. The task of allocating river solute loads among end-member sources is commonly accomplished through inversion calculations. However, the process of sampling end-member chemical values for use within river inversion models was recently made much more complex by a shift away from normalizing chemical values to sodium and toward normalizing chemical values to a sum of parameters. Given that variables normalized to a sum must themselves sum to one, the updated normalization imposes constraints on the selection of end-member chemical ratios that have previously been overlooked. Here we demonstrate that the end-member selection process employed in the initial version of the MEANDIR model, which is an open-source program for conducting river inversion calculations, does not optimally sample end-member chemistry and results in artificially high end-member chemical variances that generate excess variance in inversion results. To correct this shortcoming, we present a new variance-targeting algorithm for sum-constrained sampling of Gaussian chemical distributions. The new approach is first demonstrated for normal distributions with a constant variance, then for three normal distributions with arbitrary variances, and finally a general case is presented for an arbitrary number of normal distributions with arbitrary variances. The final case operates by numerically calculating matrix components to maximize the absolute value of the matrix determinant subject to variance-based optimization constraints; rather than employing a user-defined covariance matrix, the algorithm maximizes the joint entropy of the vector of end-member chemical ratios. The algorithm is implemented in MEANDIR and used to invert a global river chemistry dataset. The updated sampling strategy reduces the compositional range of the end-member priors relative to the initial approach and results in modest-to-large reductions in the 95th-5th percentile range of accepted results for the fraction of carbonate mineral weathering. In contrast, the median inversion result does not change substantially under the new sampling algorithm. Overall, the variance-targeting and sum-constrained sampling algorithm improves the MEANDIR model by bringing end-member chemical priors into closer alignment with user-supplied distributions.



42 1. Introduction

43 The partial pressure of carbon dioxide in Earth's atmosphere ($p\text{CO}_2$) is maintained by a
44 delicate balance between carbon and alkalinity inputs and outputs to the ocean-atmosphere system
45 (Berner and Caldeira, 1997; Kemeny et al., 2024). Whereas volcanic or metamorphic degassing,
46 oxidation of organic carbon, and oxidation of sulfide minerals increase $p\text{CO}_2$ by adding carbon or
47 removing alkalinity from the ocean-atmosphere system (Spence and Telmer, 2005; Zondervan et
48 al., 2023), processes such as burial of organic carbon and dissolution of silicate or carbonate
49 minerals reduce $p\text{CO}_2$ through removal of carbon or addition of alkalinity to the ocean-atmosphere
50 system (Berner et al., 1983; West et al., 2005; Hilton and West, 2020). Because the chemical
51 products of mineral dissolution or oxidation enter river water, where they are ultimately
52 transported to the global ocean, the chemical composition of rivers has been extensively studied
53 to constrain the dependence of weathering-derived carbon and alkalinity fluxes on local and
54 regional climatic and geomorphic conditions (Stallard and Edmond, 1983; Calmels et al., 2007;
55 Moon et al., 2014).

56 Quantitatively interpreting fluvial solute chemistry requires allocating the fraction of each
57 component in river water among possible sources. For example, calcium (Ca^{2+}) in river water could
58 potentially derive from silicate minerals, carbonate minerals, evaporite minerals, hydrothermal
59 fluids, meteoritic precipitation, or other end-member sources specific to a given landscape. To
60 address this uncertainty, river geochemists employ inversion models to unmix measured
61 parameters and constrain the distributions of possible solute contributions from end-members that
62 themselves have chemical uncertainties (Négrel et al., 1993). The inversion approach to the study
63 of river chemistry, popularized in Gaillardet et al. (1999), was summarized and extended within



64 the Mixing Elements AND Dissolved Isotopes in Rivers (MEANDIR) inversion model (Kemeny
65 and Torres, 2021).

66 The absolute concentrations of elements in river water are of limited use for constraining
67 end-member contributions because rivers can gain or lose water without gaining or losing solutes,
68 meaning that concentrations can be modified independently from the fractional contributions of
69 each end-member to each solute. As a result, the elemental ratios of river water are typically
70 employed to determine solute sources within inversion models. In a system with observations of
71 Ca^{2+} and sodium (Na^+) thought to derive from carbonate and silicate minerals, a high $\text{Ca}^{2+}/\text{Na}^+$
72 ratio would potentially suggest dominance of carbonate mineral weathering and a low $\text{Ca}^{2+}/\text{Na}^+$
73 ratio would potentially suggest dominance of silicate mineral weathering.

74 Although Na^+ was commonly used as the normalizing variable in early studies (Négrel et
75 al., 1993; Gaillardet et al., 1999), more recent work has normalized observations to a sum of
76 variables (Torres et al., 2016; Hemingway et al., 2020). Chemical observations in river water might
77 be normalized to the sum of Ca^{2+} , Na^+ , and magnesium (Mg^{2+} ; sum indicated as Σ^+) such that the
78 ratios being manipulated during the calculation are Ca^{2+}/Σ^+ , Mg^{2+}/Σ^+ , and Na^+/Σ^+ . Kemeny and
79 Torres (2021) argued that there are substantial advantages to normalizing observations to the sum
80 of dissolved parameters and extended the normalizing sum to include variables such as potassium
81 (K^+), lithium (Li^+), sulfate (SO_4^{2-}), and nitrate (NO_3^-). Normalizing to a sum removes the need for
82 all end-members to contain a common element and allows for end-members that only source one
83 element to be independently included in the inversion. Normalizing to the sum of Ca^{2+} , Mg^{2+} , Na^+ ,
84 and SO_4^{2-} would allow a carbonate end-member to not contain Na^+ while allowing for sulfide
85 oxidation to be represented as an independent end-member.



86 Despite substantial benefits, normalizing chemical ratios to a sum introduces new
87 complications into the sampling of end-member values. In particular, sum-normalization requires
88 that end-member chemical distributions be sampled such that each end-member is internally
89 consistent. When normalizing to the sum of Ca^{2+} , Mg^{2+} , and Na^{+} , consistency would require
90 sampling the $\text{Ca}^{2+}/\Sigma^{+}$, $\text{Mg}^{2+}/\Sigma^{+}$, and Na^{+}/Σ^{+} distributions so $\text{Ca}^{2+}/\Sigma^{+} + \text{Mg}^{2+}/\Sigma^{+} + \text{Na}^{+}/\Sigma^{+} = 1$.
91 Although the method of addressing this problem in the initial version of MEANDIR ensured
92 internal consistency of the end-members, it did not recreate user-supplied variances for the end-
93 member distributions (Kemeny and Torres, 2021).

94 The end-member distributions used in river inversion models are often defined as
95 Gaussians. This convention reflects both historical precedent (Gaillardet et al., 1999; Burke et al.,
96 2018) and the convenience of using a simple distribution when the form of the underlying data is
97 only loosely constrained. However, employing Gaussian distributions to describe sum-normalized
98 concentration ratios is problematic because normal distributions contain unphysical values below
99 0 and above 1. Practically, this means that the mean and variance of Gaussian distributions will be
100 modified during sampling to an extent dependent upon the fraction of probability density contained
101 within unphysical regions. Here we accept that limitation in order to prioritize the convenience of
102 working with normal distributions; although non-Gaussian distributions may better represent
103 reality, our goal in this work is to develop inversion approaches that address standard practices.
104 Due to the need for truncation, we describe the updated Gaussian sampling algorithm presented as
105 “variance-targeting” rather than “variance-preserving”. Lastly, while the problem of unphysical
106 probability density is also present in the original MEANDIR sampling algorithm, it is important
107 to note that the issue is in addition to the more fundamental problem referenced above where the



original routine fails to recreate user-supplied variances even when end-member distributions are nearly entirely bounded between 0 and 1.

110

111 **2. Sampling algorithms**

112 **2.1. Approaches that do not recreate the expected variance of end-member distributions**

113 **2.1.1. The n-1 sampling algorithm does not preserve variance**

114 The sampling algorithm of the original MEANDIR model was to sample all but one
 115 chemical distribution for each end-member and calculate the final value by closure to ensure that
 116 the sum of the normalized variables was equal to 1 (Kemeny and Torres, 2021). We call this routine
 117 the n-1 algorithm as it entails sampling n-1 of the end-member distributions independently at
 118 random before calculating the nth value by difference. If the nth value was computed to be negative,
 119 and was not supposed to be negative, MEANDIR performed the end-member selection again until
 120 it reached either a set of positive end-member chemical ratios or made a predetermined number of
 121 attempts.

122 The shortcomings of the n-1 approach readily appear in an example where end-member
 123 chemistry is normalized to the sum of Ca²⁺, Mg²⁺, and Na⁺. If Ca²⁺/Σ⁺ and Mg²⁺/Σ⁺ are pulled from
 124 normal distributions with given variances, the distribution of calculated Na⁺/Σ⁺ ratios will have
 125 $\sigma_{\text{Na}^+/\Sigma^+, n-1}^2 = \sigma_{\text{Ca}^{2+}/\Sigma^+}^2 + \sigma_{\text{Mg}^{2+}/\Sigma^+}^2$ instead of the $\sigma_{\text{Na}^+/\Sigma^+}^2$ value supplied as an input by the user.
 126 For the case of $\sigma_{\text{Ca}^{2+}/\Sigma^+} = 0.05$ and $\sigma_{\text{Mg}^{2+}/\Sigma^+} = 0.15$, $\sigma_{\text{Na}^+/\Sigma^+, n-1} \approx 0.16$ regardless of the input
 127 $\sigma_{\text{Na}^+/\Sigma^+}^2$ value (Fig. 1a). In detail, $\sigma_{\text{Na}^+/\Sigma^+, n-1}$ can also have lower variance than predicted from
 128 the addition of two Gaussian distributions sampled independently in cases where the Ca²⁺/Σ⁺ or
 129 Mg²⁺/Σ⁺ distributions have substantial probability density below 0 or above 1.

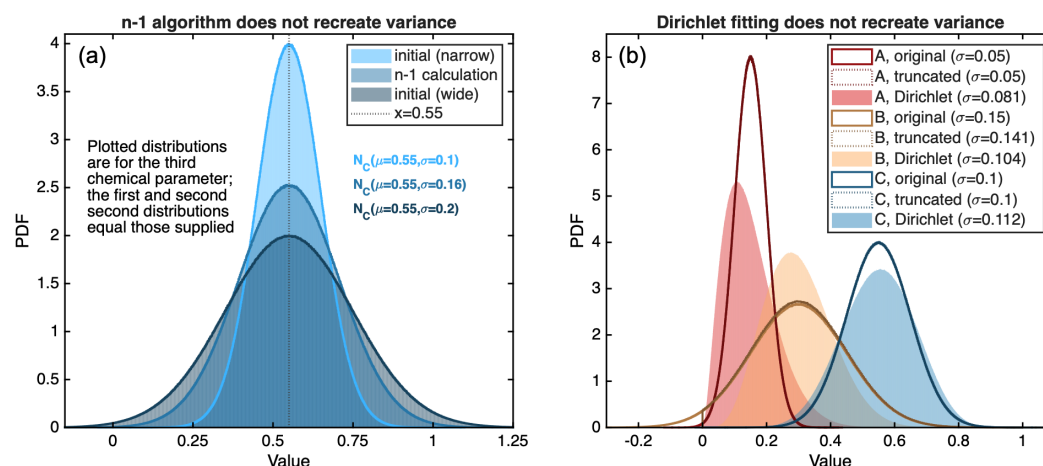


Fig. 1: Calculating end-member chemistry with the n-1 sampling strategy and with Dirichlet distributions. Data were calculated from three initial distributions: $N_A(\mu_A = 0.15, \sigma_A = 0.05)$, $N_B(\mu_B = 0.30, \sigma_B = 0.15)$, and $N_C(\mu_C = 0.55, \sigma_C = 0.10 \text{ or } 0.20)$. (a) Variance in the third chemical ratio, when calculated using the n-1 sampling algorithm, is controlled by the variance in the chemical parameters used for differencing ($\sigma_A^2 + \sigma_B^2$) and can be either larger or smaller than σ_C^2 . Given that the n-1 routine is the standard sampling approach in MEANDIR (Kemeny and Torres, 2021), it is desirable to implement improved methods for sampling Gaussian end-member chemical distributions. (b) Fitting the Dirichlet distribution to Gaussian distributions. The best-fitting Dirichlet curves, found by optimizing the α_i values to minimize misfit from the variances of truncated Gaussians while matching their means, both overestimates (red) and underestimates (yellow) target distribution variances. While the use of Dirichlet distributions is potentially appropriate when fitting actual observations, their use is not ideal for sum-constrained sampling of end-members described using Gaussian distributions.

The mismatch between the standard deviation of sampled Na^+/Σ^+ values and the user-supplied value occurs because no information about the input $\sigma_{\text{Na}^+/\Sigma^+}^2$ value entered the calculation. This example demonstrates that, prior to any consideration of truncation, the n-1 approach in the initial MEANDIR model does not preserve variances for the chemical ratio calculated by difference. Instead, the variance of the n^{th} chemical parameter is approximately equal to the sum of the variances of the other distributions, which can either be larger or smaller than the user-supplied value. A central goal of this work is thus the development and implementation of a new Gaussian sampling algorithm in which the variance of all user-supplied distributions is recreated as accurately as possible.



154 2.1.2. Fitting Dirichlet distributions to Gaussians does not preserve variance

155 One possible approach to calculating chemical distributions with the input variances is
156 through the use of Dirichlet distributions. Also known as the multivariate beta distribution, these
157 are a set of curves bounded between 0 and 1 that can be used to quantify the probability
158 distributions of values that sum to 1. In the context of river geochemistry with Ca^{2+} , Mg^{2+} , and Na^{+}
159 in the normalization, drawing a vector of probabilities from a Dirichlet distribution of order 3
160 would generate a set of internally consistent $\text{Ca}^{2+}/\Sigma^{+}$, $\text{Mg}^{2+}/\Sigma^{+}$, and Na^{+}/Σ^{+} compositions for a
161 given end-member. Unfortunately, as with the n-1 approach, the variances of the end-member
162 chemical distributions do not necessarily match the user-supplied values (Fig. 1b). The mismatch
163 of variance reflects that the Dirichlet distribution of order n has only n free parameters, typically
164 called $\alpha_1 \dots \alpha_n$. The relative magnitude of these parameters ($\alpha_i/\Sigma_j \alpha_j$) determine the position of
165 each Dirichlet curve and the absolute values of the α_i influence the distribution variances. When
166 setting the $\alpha_i/\Sigma_j \alpha_j$ values to equal the μ_i values of the end-member chemical distributions, the only
167 method for modifying the Dirichlet curves is multiplying all α values by a scalar. Although this
168 scalar can be optimized to best fit the user-supplied variances (Fig. 1b), such an optimization does
169 not force the standard deviations of the Dirichlet curves to match those of the given Gaussians.

170 Fitting and sampling Dirichlet curves results in end-member distributions where most end-
171 members have variances offset from expectations. This behavior is in contrast to the n-1 approach,
172 where n-1 of the end-members have variances closely matching expectations with the n^{th} end-
173 member potentially very different from expectations. While employing Dirichlet distributions
174 could certainly be appropriate if the underlying data follow those distributions, it is not an optimal
175 solution when end-member uncertainty is presented in the form of normal distributions. As a result,



176 the primary algorithm presented below for variance-targeting and sum-constrained Gaussian
177 sampling does not employ Dirichlet distributions.

178

179 **2.2. Variance-targeting sum-constrained algorithms for sampling chemical compositions**

180 Here we present variance-targeting sum-constrained algorithms for sampling Gaussian end-
181 member chemical distributions. We first describe sampling from distributions with the same
182 variance, then address sampling from three distributions with arbitrary variances, and finally
183 present the general case of sampling an arbitrary set of normal distributions with arbitrary
184 variances. In the final case, the solution is derived on the basis of entropy without requiring a user-
185 supplied covariance matrix.

186

187 **2.2.1. Sampling distributions with equal variance**

188 Sampling from n distributions with equal variance can be accomplished by combining
189 results from applying the $n-1$ sampling strategy to subsets of $n \cdot (n-1)$ independent pulls from the
190 standard normal distribution. This approach may seem surprising because, as demonstrated above,
191 calculating any single chemical variable as 1 minus the sum of the other variables would result in
192 the variance of the calculated distribution being equal to $\sigma^2 \cdot (n-1)$ rather than σ^2 prior to any
193 truncation. However, a solution that preserves variance can be reached by sampling symmetrically
194 across all variables and then combining those results with an appropriate normalization factor. The
195 sampling routine begins by converting each distribution into the standard normal distribution,
196 $N(\mu=0, \sigma^2=1)$, by subtracting the distribution mean and scaling by the standard deviation. The
197 problem of drawing from n distributions with distinct means and constant variance has then been
198 simplified to sampling n values, where those values have a sum of zero, and each of the n values



has the probability density function of a standard Gaussian. Once that problem is solved, the results can be scaled and translated to recreate the target distributions. The solution is first described for $n = 3$ and then generalized.

2.2.1.1. Three end-members with equal variance

The solution for $n = 3$ requires sampling six independent values from the standard normal distribution. These values can conceptually be divided into three pairs of values, where three additional values are then generated from the negative sums of the pairs. Let the six values pulled from the standard normal distribution be called A_2, A_3, B_1, B_3, C_1 , and C_2 , with the calculated values $A_1 = -(B_1 + C_1)$, $B_2 = -(A_2 + C_2)$, $C_3 = -(A_3 + B_3)$. The variance in the distribution of A_2, A_3, B_1, B_3, C_1 , and C_2 is 1, as these values were pulled from a distribution with unit variance, whereas the variance in A_1, B_2 , and C_3 is 2, as these values were calculated as the negative sum of pulls with unit variance. Using these nine values and an appropriate normalization factor, we define three new variables called A, B , and C that sum the corresponding values (eqs. 1-3). Because each term in the set of three equations appears once with a positive sign and once with a negative sign, $A + B + C = 0$. A normalizing factor of $1/2$ is required to ensure that the final A, B , and C distributions have $\sigma^2=1$. Once A, B , and C are calculated for a given set of six pulls from the standard Gaussian, the original distributions are recreated by multiplying A, B , and C by σ and then adding back the distribution means.

$$A = \frac{1}{2}(A_1 + A_2 + A_3) = \frac{1}{2}(-(B_1 + C_1) + A_2 + A_3) \quad (1)$$

$$B = \frac{1}{2}(B_1 + B_2 + B_3) = \frac{1}{2}(B_1 - (A_2 + C_2) + B_3) \quad (2)$$

$$C = \frac{1}{2}(C_1 + C_2 + C_3) = \frac{1}{2}(C_1 + C_2 - (A_3 + B_3)) \quad (3)$$



2.2.1.2. General solution for end-members with equal variance

In the solution for an arbitrary n number of variables, we again center and normalize each of the distributions to simplify the problem to pulling from the standard Gaussian. We then independently and randomly sample $n \cdot (n-1)$ values from $N(\mu=0, \sigma^2=1)$ and assign these values to the off-diagonal positions of an $n \times n$ matrix $\mathbf{M}$ with diagonal elements equal to 0 (Fig. 2). Under this organization, the value corresponding to the i^{th} distribution is the normalized sum of the entries in the i^{th} row less the sum of the entries in the i^{th} column (eq. 4). As in the simpler case, the normalizing factor of $1/\sqrt{2n-2}$ ensures that the summed distributions have $\sigma^2=1$. This normalization reflects that, for the i^{th} equation, each of $2n-2$ non-zero terms contribute a variance of 1 to the calculated values. Once the chemical ratios are calculated, the original distributions are again recreated by multiplying by the original σ and adding back the distribution means.

233

$$z_i = \frac{\sum_{j=1:n} (M_{i,j}) - \sum_{j=1:n} (M_{j,i})}{\sqrt{2n-2}} \quad (4)$$

235

236

Fig. 2: Matrix representation of a variance-targeting sum-constrained sampling algorithm for n Gaussian distributions with equal variance. For the matrix $\mathbf{M}$ containing $n \cdot (n-1)$ pulls from the standard Gaussian, where the diagonal elements are zero, the normalized i^{th} row sum minus the i^{th} column sum defines an instance of the i^{th} target distribution. Highlighted elements indicate variables used to calculate the first (red), third (blue), and n^{th} (yellow) chemical values.

253

$$z_1 = (\sum_{j=1:n} (M_{1,j}) - \sum_{j=1:n} (M_{j,1})) / \sqrt{(2n-2)}$$

$$z_3 = (\sum_{j=1:n} (M_{3,j}) - \sum_{j=1:n} (M_{j,3})) / \sqrt{(2n-2)}$$

$$z_n = (\sum_{j=1:n} (M_{n,j}) - \sum_{j=1:n} (M_{j,n})) / \sqrt{(2n-2)}$$

0	$M_{1,2}$	$M_{1,3}$	...	$M_{1,n}$
$M_{2,1}$	0	$M_{2,3}$	...	$M_{2,n}$
$M_{3,1}$	$M_{3,2}$	0	...	$M_{3,n}$
...	...	...	...	...
$M_{n,1}$	$M_{n,2}$	$M_{n,3}$	...	0



2.2.2. Sampling distributions with arbitrary variances

2.2.2.1. Three end-members with arbitrary variances

Consider the case of three Gaussians: $N_A^0 = N(\mu_A, \sigma_A^2)$, $N_B^0 = N(\mu_B, \sigma_B^2)$, $N_C^0 = N(\mu_C, \sigma_C^2)$, arranged such that $\sigma_A^2 \leq \sigma_B^2 \leq \sigma_C^2$ and where $\mu_A + \mu_B + \mu_C = 1$. As in the preceding example, center each distribution by removing their means to reach $N_A = N(\mu = 0, \sigma_A^2)$, $N_B = N(\mu = 0, \sigma_B^2)$, $N_C = N(\mu = 0, \sigma_C^2)$. Our approach to sampling these distributions is to produce two different sets of distributions, one with variance matching or larger than the target distributions (A_{high} , B_{high} , C_{high}) and one with variance matching or less than the target distributions (A_{low} , B_{low} , C_{low}), which are sampled independently and linearly combined to reach distributions with the target variances.

The maximum possible variance in values of C occurs when the pulls from distribution A are perfectly correlated with pulls from distribution B. To reach such a C distribution, sample the value A_{high} from N_A and calculate B_{high} as $A_{\text{high}} \cdot \sigma_B / \sigma_A$. The value C_{high} then equals $-(A_{\text{high}} + B_{\text{high}})$. Under this formulation, the distribution of A_{high} has variance σ_A^2 , the distribution of B_{high} has variance σ_B^2 , and the distribution of C_{high} has variance $(\sigma_A + \sigma_B)^2$. The variance of C_{high} is not $(\sigma_A^2 + \sigma_B^2)$, which would result from adding independently drawn values, but has been further increased by $2 \cdot \sigma_A \cdot \sigma_B$ through covariation between A_{high} and B_{high} .

Second, consider a set of choices that will result in a variance for the distribution C that is lower than σ_C^2 . Again, we begin by sampling N_A directly to find an individual instance of A_{low} . Second, draw an instance of C_{low} by sampling the distribution $N(\mu = 0, \sigma_B^2 - \sigma_A^2)$, where the imposed ordering guarantees that $\sigma_B^2 - \sigma_A^2$ is non-negative and less than σ_C^2 . In turn, B_{low} is set equal to $-(A_{\text{low}} + C_{\text{low}})$, which will then have variance of $\sigma_{A_{\text{low}}}^2 + \sigma_{C_{\text{low}}}^2 = \sigma_A^2 + (\sigma_B^2 - \sigma_A^2) =$



276 σ_B^2 . Overall, the distribution of A_{low} has variance σ_A^2 , the distribution of B_{low} has variance σ_B^2 , and
277 the distribution of C_{low} has variance $(\sigma_B^2 - \sigma_A^2)$.

278 The two sets of distributions are then combined to produce new distributions with the target
279 variances. To do so, define C as a linear combination of C_{high} and C_{low} , where we combine an
280 amount p of C_{high} and an amount $\sqrt{1 - p^2}$ of C_{low} such that the resulting distribution has a
281 variance of σ_C^2 (eqs. 5-7). The value p is found to depend only on σ_A , σ_B , and σ_C (eq. 8), and the
282 values of A , B , and C can be directly calculated (eqs. 9-11). The sum of A , B , and C is zero because
283 the sum of A_{high} , B_{high} , and C_{high} is zero and the sum of A_{low} , B_{low} , and C_{low} is zero. By virtue
284 of the calculation, the variances of A , B , and C , are σ_A^2 , σ_B^2 , and σ_C^2 . As in the case of constant
285 variance, the target distributions are recreated by adding μ_A , μ_B , and μ_C back to A , B , and C .

286

$$287 \quad C = p \cdot C_{high} + \sqrt{1 - p^2} \cdot C_{low} \quad (5)$$

$$288 \quad \text{Var}(C) = p^2 \cdot \text{Var}(C_{high}) + (1 - p^2) \cdot \text{Var}(C_{low}) \quad (6)$$

$$289 \quad \sigma_C^2 = p^2 \cdot (\sigma_A^2 + \sigma_B^2 + 2\sigma_A\sigma_B) + (1 - p^2) \cdot (\sigma_B^2 - \sigma_A^2) \quad (7)$$

$$290 \quad p = \sqrt{\frac{\sigma_A^2 + \sigma_C^2 - \sigma_B^2}{2\sigma_A^2 + 2\sigma_A\sigma_B}} \quad (8)$$

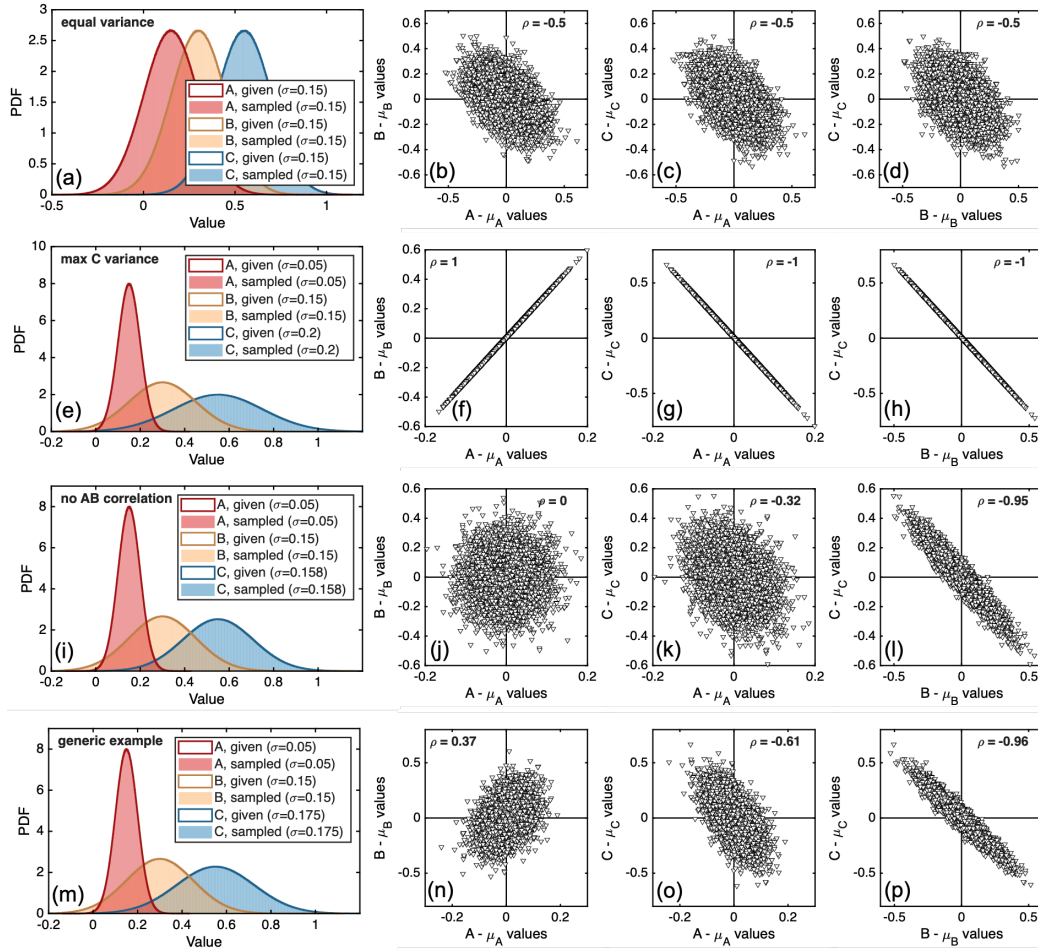
$$291 \quad A = p \cdot A_{high} + \sqrt{1 - p^2} \cdot A_{low} \quad (9)$$

$$292 \quad B = p \cdot B_{high} + \sqrt{1 - p^2} \cdot B_{low} \quad (10)$$

$$293 \quad C = p \cdot C_{high} + \sqrt{1 - p^2} \cdot C_{low} \quad (11)$$

294

295 The described sampling algorithm generates three sets of distributions that match the user-
296 supplied distributions (Fig. 3). It is important to appreciate, however, that the relationship between
297 instances of A , B , and C depends strongly on the relationships among σ_A^2 , σ_B^2 , and σ_C^2 . For example,



298

Fig. 3: Example of a variance-targeting algorithm that performs sum-constrained sampling on 3 Gaussians with positive variances ordered such that $\sigma_A^2 \leq \sigma_B^2 \leq \sigma_C^2$. Sampling is illustrated for four different sets of initial distributions (**a**, **e**, **i**, **m**) where the means of each distribution, but not the variances, are preserved across the four cases ($\mu_A = 0.15$, $\mu_B = 0.30$, $\mu_C = 0.55$). Differences among the results of the second, third, and fourth rows derive entirely from changes in σ_C , which regulates the covariances among A, B, and C. Depending on the relationships among σ_A , σ_B , and σ_C , values of B can be (**b**) negatively correlated with A, (**f**, **n**) positively correlated with A, or (**j**) uncorrelated with A. (**f**–**h**) Values of B have perfect positive correlation to A and values of C have perfect negative correlation to A when $\sigma_C = \sigma_A + \sigma_B$.

when $\sigma_A^2 = \sigma_B^2 = \sigma_C^2$, the correlations between A, B, and C are all equal to -0.5 (Figs. 3b-d). In

contrast, when the variance of C is maximized, meaning that $\sigma_C^2 = (\sigma_A + \sigma_B)^2$, values of A and B



are perfectly correlated with one another and perfectly anticorrelated with C (Figs. 3f-h). In contrast, when $\sigma_C^2 = \sigma_A^2 + \sigma_B^2$ (Fig. 3j-l), the correlation between A and B is zero. At an intermediate case, the distribution of C values is anticorrelated with values of both A and B (Figs. 3n-p), although C is correlated more strongly with B than with A. These results demonstrate that the variance of end-member distributions exerts a large influence on the covariance between end-member chemical values. Practically, this means that changes in the user-supplied variance of an individual chemical ratio will modify the degree of covariation among the other chemical variables in the analysis.

319

2.2.2.2. General solution for end-members with arbitrary variances

For the general case, consider n Gaussians with arbitrary variances where the i^{th} distribution is defined as $N_i^0 = N(\mu_i, \sigma_i^2)$ and $\Sigma_i(\mu_i) = 1$. As in the prior examples, we first subtract the mean of each distribution to generate centered distributions where $N_i = N(\mu = 0, \sigma_i^2)$. Next, to understand why the problem with $n > 3$ is meaningfully different from the problem where $n = 3$, consider that only three correlations were involved in the earlier case: values of A with those of B and C and values of B with those of C. For $n = 4$, six correlations must be considered: values of A with those of B, C, and D, values of B with those of C and D, and values of C with those of D. In general, the challenge of sampling $n > 3$ end-member chemical variables can be viewed as the problem of selecting how the sampled values should correlate with one another. There are many possible options for this correlation that satisfy the constraints that each variable's distribution has the required variance, required mean, and that the individual sets of values sum to 1, and the problem is well-expressed through an example of a poor choice of correlation. Consider that, for $n = 4$ with $\sigma_A^2 = \sigma_B^2 = \sigma_C^2 = \sigma_D^2$, we might choose to sample from $N_A(\mu = 0, \sigma_A^2)$ and then set



334 $B = A$, $C = -A$, and $D = -A$. In this case, $A + B + C + D = 0$ and the variances of A , B , C , and
 335 D are all σ_A^2 , satisfying the main criteria for our sampling. Despite meeting the criteria, this
 336 algorithm generates perfect covariance between A and B , as well as between C and D , with perfect
 337 anticorrelation between A or B with C or D . Such correlation attributes are clearly non-desirable as
 338 they impose structure on the sampled values that is not required by the problem.

339 An improved correlation scheme would, intuitively, be one in which knowledge of any one
 340 chemical variable reveals as little information as possible regarding the remaining variables.
 341 Furthermore, it is desirable for the calculated distributions to not contain any correlations stronger
 342 than necessary. Formally, we seek a sampling algorithm that maximizes the joint entropy of the
 343 vector of end-member chemical ratios. In the $n = 4$ above example, where knowing any individual
 344 value constrained the entire system, the sampling routine is problematic specifically because the
 345 entropy of the vector of end-member chemical ratios is very low.

346 To maximize the entropy of the sampling, we define the $n \times 1$ vector of chemical ratios ($\vec{z}$)
 347 for a given end-member as weighted linear combinations of $n-1$ values pulled from a standard
 348 normal distribution ($\vec{m}$) plus the $n \times 1$ vector of distribution means ($\vec{\mu}$) (Fig. 4). Let the weighting
 349 terms be indicated as $G_{i,j}$, where both i and j range from 1 to $n-1$, arranged into an $(n-1) \times (n-1)$
 350 matrix indicated as $\mathbf{G}$. Let an augmented version of $\mathbf{G}$ containing an extra row for closure be
 351 defined as the $n \times (n-1)$ matrix $\tilde{\mathbf{G}}$; under these definitions, $\vec{z} = \tilde{\mathbf{G}}\vec{m} + \vec{\mu}$ (Fig. 4). To find the $G_{i,j}$
 352 values that maximize the joint entropy of $\vec{z}$, we maximize the absolute value of the determinant of
 353 $\mathbf{G}$ while subject to optimization constraints that ensure the variance of the distribution of each $\vec{z}$
 354 value matches expectations. Optimizing $|\det(\mathbf{G})|$ maximizes the entropy of the sampling because
 355 $h(\tilde{\mathbf{G}}\vec{m}) = h(\vec{m}) + \log(|\det(\mathbf{G})|)$ for a random variable $\vec{m}$, an invertible matrix $\mathbf{G}$, and with h denoting
 356 entropy. As a result, for a fixed random variable $\vec{m}$ (as in our case, where $\vec{m}$ is defined to contain



357 n-1 entries drawn independently from the standard Gaussian), the entropy $h(\vec{G}\vec{m})$ attains its
358 maximum value when $|\det(\mathbf{G})|$ attains its maximum value.

359

$$\begin{array}{c}
 \vec{z} \\
 \left[\begin{array}{c} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_{n-1} \\ z_n \end{array} \right] = \begin{array}{c} \mathbf{G}/\tilde{\mathbf{G}} \\ \left[\begin{array}{ccccc} G_{1,1} & G_{1,2} & G_{1,3} & \dots & G_{1,n-1} \\ G_{2,1} & G_{2,2} & G_{2,3} & \dots & G_{2,n-1} \\ G_{3,1} & G_{3,2} & G_{3,3} & \dots & G_{3,n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_{n-1,1} & G_{n-1,2} & G_{n-1,3} & \dots & G_{n-1,n-1} \end{array} \right] \left[\begin{array}{c} m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_{n-1} \end{array} \right] + \left[\begin{array}{c} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_{n-1} \\ \mu_n \end{array} \right] \\
 \left[\begin{array}{ccccc} -\sum_i^{n-1}(G_{i,1}) & -\sum_i^{n-1}(G_{i,2}) & -\sum_i^{n-1}(G_{i,3}) & \dots & -\sum_i^{n-1}(G_{i,n-1}) \end{array} \right]
 \end{array}
 \end{array}$$

360

Values to be used in the inversion

Values constrained by maximizing the determinant of $\mathbf{G}$ subject to n optimization constraints that account for distribution variances

Values from $N(0,1)$

Means of the distributions

361 **Fig. 4:** Schematic of the algorithm for variance-targeting sum-constrained sampling of an arbitrary
362 number of Gaussians with arbitrary variances. For n distributions, the $(n-1) \times (n-1)$ matrix $\mathbf{G}$
363 (within the solid bracket) is optimized to maximize the absolute value of the determinant of $\mathbf{G}$
364 while meeting n optimization criteria ensuring that all sampled distributions have the expected
365 variances (eqs. 12, 13). Once the $G_{i,j}$ elements have been constrained, the $n \times (n-1)$ matrix $\tilde{\mathbf{G}}$ (within
366 the dotted extension of the bracket) is calculated by closure. Multiplication of $\tilde{\mathbf{G}}$ by a vector of n-
367 1 values pulled from a standard normal distribution, followed by addition of the distribution means,
368 generates a vector of n values for use in inversion calculations. The optimization criteria can also
369 be adjusted to account for additional compositional constraints on the evaporite end-member.
370 Multiple $\mathbf{G}$ matrices related through high-dimensional rotations may satisfy the optimization
371 criteria with equal determinants and any such matrix can be used to generate $\tilde{\mathbf{G}}$ and $\vec{z}$.
372

373 Optimization constraints ensure that the variance of the $\vec{z}$ values equals the user-supplied
374 variances. Out of n constraints, n-1 entail setting the sum of the squared $G_{i,j}$ values along each row
375 equal to the target variance for that entry (eq. 12). The n^{th} optimization criterion requires that the
376 sum of the squares of the column sums of $\mathbf{G}$ equals the variance of the final distribution (eq. 13).



377 This last constraint reflects that the final chemical variable will be calculated as the negative sum
 378 of the other values (Fig. 4).

379

$$380 \quad \sigma_{i \ (i < n)}^2 = \sum_{j=1}^{n-1} (G_{ij}^2) \quad (12)$$

$$381 \quad \sigma_n^2 = \sum_{j=1}^{n-1} \left(\sum_{i=1}^{n-1} (G_{ij}) \right)^2 \quad (13)$$

382

383 Unlike other Gaussian sampling schemes, the new entropy-maximizing approach does not
 384 require specifying a covariance matrix (Altmann et al., 2014). Rather, covariance between the
 385 variables is generated implicitly through the optimization of **G**. While the proposed agnostic
 386 approach works well when the covariance structure among chemical values is unknown, it does
 387 not present the best option when certain variables are known to be particularly correlated or
 388 anticorrelated with one another.

389 For a specific example of the generalized sampling algorithm, consider the case of $n = 6$
 390 with $\sigma_A = \sigma_B = \sigma_C = \sigma_D = \sigma_E = \sigma_F = 0.1$. The sampling algorithm for this case requires
 391 optimizing the 25 elements of a 5×5 matrix to maximize the absolute value of its determinant
 392 subject to six optimization constraints (eqs. 14-19). Because a family of matrices related through
 393 high-dimensional rotations can all satisfy the optimization criteria at the same determinant value
 394 (in this case, 6.44×10^{-6}), the exact matrix elements calculated through optimization will depend
 395 on the specific initial conditions of the optimization. Indeed, many different but interchangeable
 396 **G** matrices can be found (eq. 20 and eq. 21). Despite having different matrix elements, the two
 397 matrices will generate viable distributions for end-member chemistry. That is, even though there
 398 exist multiple matrices **G** that maximize the determinant subject to the same variance-based
 399 optimization constraints, they are generally all related by rotations and the output distributions



they lead to are identical. An instance of $\mathbf{G}$ for a more complex case is also provided where the normalization includes Ca^{2+} , Mg^{2+} , Na^+ , K^+ , Sr^{2+} , and Li^+ with $\sigma_{\text{Ca}^{2+}/\Sigma^+} = 0.20$, $\sigma_{\text{Mg}^{2+}/\Sigma^+} = 0.15$, $\sigma_{\text{Na}^+/\Sigma^+} = 0.10$, $\sigma_{\text{K}^+/\Sigma^+} = 0.05$, $\sigma_{\text{Sr}^{2+}/\Sigma^+} = 5 \times 10^{-4}$, and $\sigma_{\text{Li}^+/\Sigma^+} = 5 \times 10^{-6}$ (eqs. 14-19, 22). As in this example, the largest standard deviation of the user-supplied Gaussian distributions should generally be less than the sum of the other standard deviations.

$$\sigma_A^2 \text{ or } \sigma_{\text{Li}^+/\Sigma^+}^2 = G_{1,1}^2 + G_{1,2}^2 + G_{1,3}^2 + G_{1,4}^2 + G_{1,5}^2 \quad (14)$$

$$\sigma_B^2 \text{ or } \sigma_{\text{Sr}^{2+}/\Sigma^+}^2 = G_{2,1}^2 + G_{2,2}^2 + G_{2,3}^2 + G_{2,4}^2 + G_{2,5}^2 \quad (15)$$

$$\sigma_C^2 \text{ or } \sigma_{\text{K}^+/\Sigma^+}^2 = G_{3,1}^2 + G_{3,2}^2 + G_{3,3}^2 + G_{3,4}^2 + G_{3,5}^2 \quad (16)$$

$$\sigma_D^2 \text{ or } \sigma_{\text{Na}^+/\Sigma^+}^2 = G_{4,1}^2 + G_{4,2}^2 + G_{4,3}^2 + G_{4,4}^2 + G_{4,5}^2 \quad (17)$$

$$\sigma_E^2 \text{ or } \sigma_{\text{Mg}^{2+}/\Sigma^+}^2 = G_{5,1}^2 + G_{5,2}^2 + G_{5,3}^2 + G_{5,4}^2 + G_{5,5}^2 \quad (18)$$

$$\begin{aligned} \sigma_F^2 \text{ or } \sigma_{\text{Ca}^{2+}/\Sigma^+}^2 = & (G_{1,1} + G_{2,1} + G_{3,1} + G_{4,1} + G_{5,1})^2 + (G_{1,2} + G_{2,2} + G_{3,2} + G_{4,2} + G_{5,2})^2 + \\ & (G_{1,3} + G_{2,3} + G_{3,3} + G_{4,3} + G_{5,3})^2 + (G_{1,4} + G_{2,4} + G_{3,4} + G_{4,4} + G_{5,4})^2 + (G_{1,5} + G_{2,5} + \\ & G_{3,5} + G_{4,5} + G_{5,5})^2 \end{aligned} \quad (19)$$

$$\mathbf{G}_{\sigma_i=0.1}^{\text{version 1}} = \begin{bmatrix} -0.0422 & -0.0083 & +0.0129 & +0.0543 & -0.0710 \\ -0.0188 & +0.0139 & -0.0324 & +0.0546 & +0.0736 \\ +0.0552 & -0.0700 & -0.0428 & -0.0096 & -0.0116 \\ +0.0681 & +0.0658 & +0.0295 & +0.0036 & -0.0119 \\ -0.0419 & +0.0347 & -0.0450 & -0.0697 & -0.0125 \end{bmatrix} \quad (20)$$

$$\mathbf{G}_{\sigma_i=0.1}^{\text{version 2}} = \begin{bmatrix} +0.0966 & -0.0130 & -0.0130 & -0.0130 & -0.0130 \\ -0.0130 & +0.0966 & -0.0130 & -0.0130 & -0.0130 \\ -0.0130 & -0.0130 & +0.0966 & -0.0130 & -0.0130 \\ -0.0130 & -0.0130 & -0.0130 & +0.0966 & -0.0130 \\ -0.0130 & -0.0130 & -0.0130 & -0.0130 & +0.0966 \end{bmatrix} \quad (21)$$



$$\mathbf{G}_{\sigma \text{ variable}} = \begin{bmatrix} +0.000001 & +0.000004 & -0.000002 & -0.000001 & +0.000002 \\ -0.000286 & +0.000128 & +0.000109 & -0.000335 & -0.000167 \\ -0.022998 & +0.015721 & -0.016068 & +0.031320 & -0.022018 \\ -0.057370 & -0.036716 & -0.052332 & -0.007533 & +0.050649 \\ -0.061164 & +0.044490 & +0.079793 & +0.051060 & +0.088349 \end{bmatrix} \quad (22)$$

417

418 Within the river inversion context, each end-member included in the calculation defines a
 419 $\mathbf{G}$ matrix that can be calculated before Monte Carlo sampling begins. The requisite constrained
 420 optimization of the matrix determinant can be completed efficiently (~seconds) within MATLAB.
 421 Once each $\mathbf{G}$ matrix is optimized for the given set of σ_i^2 values, individual instances of end-
 422 member chemistry can be generated rapidly as they only require multiplying $\tilde{\mathbf{G}}$ by $n-1$ values pulled
 423 from the standard Gaussian and adding the distribution means (Fig. 4).

424 The preceding algorithm generates vectors of end-member chemistry that preserve user-
 425 supplied variances and sum to one without requiring a user-supplied covariance matrix (Altmann
 426 et al., 2014). However, not all vectors generated through the sampling algorithm will be physically
 427 realizable as some components may attain negative values. When this occurs, MEANDIR repeats
 428 the sampling process until either the end-member chemistry vector is entirely positive or a
 429 maximum number of attempts has been reached. As alluded to above, the process of rejecting
 430 negative results essentially truncates the distributions and modifies their moments relative to user-
 431 supplied values, converting a variance-preserving sampling algorithm into a routine that is merely
 432 variance-targeting. Although a reduction in variance is a natural outcome when using sum-
 433 normalized normal distributions to define variables that must take on values between 0 and 1, not
 434 all choices of Gaussian distribution are equally affected. In general, user-supplied distributions
 435 with minimal probability density below 0 or above 1 will be more faithfully recreated than
 436 distributions with density outside the physically-allowable window. For example, $N(\mu_i = 0.3, \sigma_i =$



0.1) fits well into the range of attainable chemical ratios while $N(\mu_i = 0.8, \sigma_i = 0.5)$ does not, and the moments of the latter distribution will be modified during the sampling algorithm as a non-trivial portion of the distribution would be truncated.

440

2.2.2.3. The evaporite end-member modifies the optimization constraints

The evaporite end-member in river inversion calculations is often assumed to be a stoichiometric balance of anhydrite (CaSO_4), epsomite (MgSO_4), celestine (SrSO_4), halite (NaCl), sylvite (KCl), and other metal chlorides (such as LiCl). In cases where SO_4^{2-} and Cl^- are not included in the normalization, the compositional constraint on the evaporite can be applied by setting its $\text{SO}_4^{2-}/\Sigma^+$ ratio equal to the sum of its Ca^{2+}/Σ^+ , Mg^{2+}/Σ^+ , and Sr^{2+}/Σ^+ ratios and by setting its Cl^-/Σ^+ ratio equal to the sum of its Na^+/Σ^+ , K^+/Σ^+ , Li^+/Σ^+ , and other cation ratios. This can be done when employing either the n-1 or variance-targeting sampling algorithms. Under the n-1 sampling algorithm, it remains relatively simple to account for compositional constraints when either SO_4^{2-} or Cl^- are included in the normalization. In this case, the stoichiometric relationships are attained by simultaneously solving for the closure variable and the anion ratio that together achieve the desired chemical relationships (Kemeny and Torres, 2021). Here we describe how to impose the same evaporite constraint under the variance-targeting algorithm when SO_4^{2-} is included in the normalization by modifying the optimization constraints of **G**.

Consider a scenario where river observations and end-members are normalized to the sum of Ca^{2+} , Mg^{2+} , Na^+ , and SO_4^{2-} (indicated as $\Sigma^\pm$), and where $\text{SO}_4^{2-}/\Sigma^\pm = \text{Ca}^{2+}/\Sigma^\pm + \text{Mg}^{2+}/\Sigma^\pm$. By defining $\text{SO}_4^{2-}/\Sigma^\pm$ as a function of the other variables, the dimensionality of the optimization has been reduced by 1 and $\sigma_{\text{SO}_4^{2-}/\Sigma^\pm}^2$ is no longer a free parameter of the inversion. In this specific case, the $\text{SO}_4^{2-}/\Sigma^\pm$ ratio is equal to $\frac{1}{2}$ of 1 minus the normalized $\text{Na}^+/\Sigma^\pm$ ratio (eq. 23). Given the



relationships among $\text{Na}^+/\Sigma^\pm$, $\text{Ca}^{2+}/\Sigma^\pm$, and $\text{Mg}^{2+}/\Sigma^\pm$, the standard deviation of the $\text{Na}^+/\Sigma^\pm$ ratio is also constrained such that $2|\sigma_{\text{Ca}} - \sigma_{\text{Mg}}| < \sigma_{\text{Na}} < 2(\sigma_{\text{Ca}} + \sigma_{\text{Mg}})$. Given that $\text{SO}_4^{2-}/\Sigma^\pm$ contributes to the calculation of the ion solved by closure but is no longer explicitly represented within the $\mathbf{G}$ matrix, the optimization constraint for the closure variable must be adjusted relative to the standard constraints (eqs. 13, 19). In particular, the sampling approach for this example would be to maximize $\left| \det \begin{pmatrix} G_{1,1} & G_{1,2} \\ G_{2,1} & G_{2,2} \end{pmatrix} \right|$ subject to $n-1$ optimization constraints (eqs. 24-26, note the $\frac{1}{2}$ factors on $G_{1,1}$ and $G_{1,2}$ in eq. 26). Once the $(n-2) \times (n-2)$ $\mathbf{G}$ matrix is optimized, the evaporite ratios for use in the inversion can be quickly calculated from the augmented matrix (eq. 27).

$$\frac{\text{SO}_4^{2-}}{\Sigma^\pm} = \left(\frac{\text{Ca}^{2+}}{\Sigma^\pm} + \frac{\text{Mg}^{2+}}{\Sigma^\pm} \right) = \frac{1}{2} \left(1 - \frac{\text{Na}^+}{\Sigma^\pm} \right) \quad (23)$$

$$\sigma_{\text{Na}^+/\Sigma^\pm}^2 = G_{1,1}^2 + G_{1,2}^2 \quad (24)$$

$$\sigma_{\text{Mg}^{2+}/\Sigma^\pm}^2 = G_{2,1}^2 + G_{2,2}^2 \quad (25)$$

$$\sigma_{\text{Ca}^{2+}/\Sigma^\pm}^2 = \left(\frac{1}{2} G_{1,1} + G_{2,1} \right)^2 + \left(\frac{1}{2} G_{1,2} + G_{2,2} \right)^2 \quad (26)$$

$$\begin{bmatrix} \text{Na}^+/\Sigma^\pm \\ \text{Mg}^{2+}/\Sigma^\pm \\ \text{SO}_4^{2-}/\Sigma^\pm \\ \text{Ca}^{2+}/\Sigma^\pm \end{bmatrix} = \begin{bmatrix} G_{1,1} & G_{1,2} \\ G_{2,1} & G_{2,2} \\ -\frac{1}{2} G_{1,1} & -\frac{1}{2} G_{1,2} \\ -\left(\frac{1}{2} G_{1,1} + G_{2,1}\right) & -\left(\frac{1}{2} G_{1,2} + G_{2,2}\right) \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} + \begin{bmatrix} \mu_{\text{Na}^+/\Sigma^\pm} \\ \mu_{\text{Mg}^{2+}/\Sigma^\pm} \\ \mu_{\text{SO}_4^{2-}/\Sigma^\pm} \\ \mu_{\text{Ca}^{2+}/\Sigma^\pm} \end{bmatrix} \quad (27)$$

The above example demonstrates a specific instance of balancing the evaporite end-member where SO_4^{2-} is included in the normalization. While this approach can be readily adapted conceptually to account for additional variables in the normalization, the code associated with the updated sampling algorithm does not include a pre-established method for balancing the evaporite mineral across all possible inversion designs. That is, because the function for maximizing the



determinant of $\mathbf{G}$ required hard-coding optimization constraints, the code has little built-in flexibility for modifying the chemical variables included in an evaporite end-member subject to compositional constraints. Instead, the model must be updated to reflect changes to the evaporite stoichiometry. For the inversion results presented in the supplement of this article (Figs. S1-S3), where SO_4^{2-} is included in the normalization variable, the routine described above is used to constrain the evaporite end-member such that $\text{SO}_4^{2-}/\Sigma^{\pm} = \text{Ca}^{2+}/\Sigma^{\pm} + \text{Mg}^{2+}/\Sigma^{\pm}$ (eqs. 23-27). The $\text{Cl}^-/\Sigma^{\pm}$ ratio is simply set equal to the $\text{Na}^+/\Sigma^{\pm}$ ratio.

3. Numerical implementation of the new sampling algorithm in MEANDIR

The variance-targeting and sum-constrained sampling algorithm for Gaussian distributions was implemented in the end-member selection module of the MEANDIR river inversion model (Kemeny and Torres, 2021). This implementation required several modifications to the baseline mathematics to become numerically tractable within MATLAB. Notable details of the numerical implementation include the specific choice of optimization cost function, the definition of the optimization variables, and the ordering of the chemical variables.

Optimizing $|\det(\mathbf{G})|$ is a numerically difficult problem for the *fmincon* function within MATLAB. In part, the difficulty arises because $|\det(\mathbf{G})|$ can change dramatically as the entries of $\mathbf{G}$ are modified. The issue can be partly mitigated by optimizing for $\log(|\det(\mathbf{G})|)$ rather than $|\det(\mathbf{G})|$, and this change is acceptable because $\log(|\det(\mathbf{G})|)$ attains its maximum value when $|\det(\mathbf{G})|$ is also maximized. Second, we convert $\det(\mathbf{G})$ into a more stable form by performing a singular value decomposition (eq. 28), where $\mathbf{U}$ and $\mathbf{V}^T$ are orthogonal matrices and $\mathbf{S}$ is a diagonal matrix with real non-negative values. Since $|\det(\mathbf{U})| = |\det(\mathbf{V}^T)| = 1$ as $\mathbf{U}$ and $\mathbf{V}^T$ are orthogonal and $|\det(\mathbf{S})| = \det(\mathbf{S})$ because the singular values are non-negative, $|\det(\mathbf{G})|$ simplifies to the product of



the singular values (eq. 29). In turn, $\log(|\det(\mathbf{G})|)$ simplifies to the sum of the logarithms of the singular values (eq. 30). This formulation of the optimization adds numerical stability as it avoids ever taking the product of a large number of variables, and it is this final form of the cost function that is evaluated in the updated version of MEANDIR.

507

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (28)$$

$$|\det(\mathbf{G})| = |\det(\mathbf{U}\mathbf{S}\mathbf{V}^T)| = |\det(\mathbf{U})||\det(\mathbf{S})||\det(\mathbf{V}^T)| = \det(\mathbf{S}) = \prod_{i=1}^{n-1} (S_{i,i}) \quad (29)$$

$$\log(|\det(\mathbf{G})|) = \log\left(\prod_{i=1}^{n-1} (S_{i,i})\right) = \sum_{i=1}^{n-1} \log(S_{i,i}) \quad (30)$$

511

The second complication with the numerical optimization is that, when the equations are written with the optimization constraints described above (eqs. 12, 13), the $G_{i,j}$ variables are potentially required to attain extremely different magnitudes depending on the variances of the target distributions. To normalize the scale of the matrix entries, we instead define $\mathbf{G}$ as the product of an $(n-1) \times (n-1)$ diagonal matrix $\mathbf{D}$ and an $(n-1) \times (n-1)$ matrix $\mathbf{X}$, where the diagonal entries of $\mathbf{D}$ are equal to the first $n-1$ standard deviations of the target distributions (eq. 31). Under this formulation, we maximize the cost function evaluated on $\mathbf{D}\mathbf{X}$ subject to modified optimization constraints operating on the entries of $\mathbf{X}$. The first $n-1$ modified constraints state that the sum of the squared values in each row of $\mathbf{X}$ must equal 1 (eq. 32). The n^{th} constraint is that the squared sum of the column entries, weighted by the relevant standard deviations, must equal the variance of the n^{th} chemical distribution (eq. 33). Each entry of $\mathbf{X}$ is bounded between -1 and 1, and the initial condition for each $X_{i,j}$ is sampled randomly and independently from a uniform distribution between -1 and 1.

525



$$\mathbf{G} = \mathbf{D}\mathbf{X} \quad (31)$$

$$1 = \sum_{j=1}^{n-1} (X_{i,j}^2) \text{ (for } i \leq n-1) \quad (32)$$

$$\sigma_n^2 = \sum_{j=1}^{n-1} \left(\sum_{i=1}^{n-1} (X_{i,j} \sigma_i) \right)^2 \quad (33)$$

529

530 Lastly, we observed that the optimization function works more efficiently when the final
 531 standard deviation, σ_n , is relatively large. As a result, the input list of standard deviations is sorted
 532 from smallest to largest prior to the optimization of $\mathbf{G}$. Furthermore, chemical variables with
 533 standard deviations of zero are removed prior to the calculation. The rows of $\mathbf{G}$ are then rearranged
 534 after the optimization is complete and rows containing zeros are added to the matrix in order to
 535 match the user-prescribed ordering of the chemical variables.

536

537 **4. Results and discussion**

538 The n-1 algorithm and the variance-targeting algorithm were used in inversions of the global
 539 river dataset of Gaillardet et al. (1999) to evaluate how changes in end-member sampling modify
 540 inversion outcomes. The results discussed below included Ca^{2+} , Mg^{2+} , Na^+ , and Cl^- observations
 541 in the inversions. Both inversions used the same set of end-members representing silicate minerals,
 542 carbonate minerals, evaporite minerals, and meteoric precipitation, with chemical compositions
 543 normalized to the sum of Ca^{2+} , Mg^{2+} , and Na^+ (Table 1). The n-1 inversion employed Ca^{2+} as the
 544 closure variable for the carbonate end-member and Na^+ as the closure variable for the silicate,
 545 evaporite, and precipitation end-members. The inversions did not use Cl^- critical values, the
 546 inversion cost functions were proportional to the misfit with equal variable weightings, and
 547 individual Monte Carlo simulations were accepted if they reproduced between 85% and 115% of
 548 each river observation and the normalization variable (Kemeny and Torres, 2021). The model was



run on the high-performance computing cluster at the Massachusetts Institute of Technology, and
 most samples (58/62) were successfully inverted 2,000 times.

Table 1: Gaussian end-member chemical distributions for carbonate, silicate, evaporite, precipitation, and sulfide oxidation end-members. Values are normalized to either the sum of Ca^{2+} , Mg^{2+} , and Na^+ (Σ^+) or Ca^{2+} , Mg^{2+} , Na^+ , and SO_4^{2-} ($\Sigma^\pm$) in units of charge equivalents. For the n-1 inversion results, Ca^{2+}/Σ^+ is calculated by closure in the carbonate end-member and Na^+/Σ^+ is calculated by closure in the silicate, evaporite, and precipitation end-members. Given values are means ± 1 standard deviation.

Inversion end-member chemical distributions					
End-member	Ca^{2+}	Mg^{2+}	Na^+	Cl^-	SO_4^{2-}
Normalized to $\text{Ca}^{2+} + \text{Mg}^{2+} + \text{Na}^+$ (Σ^+), results shown in Figures 5-7					
Carbonate	0.660 ± 0.084	0.340 ± 0.084	0.000 ± 0.000	0.000 ± 0.000	-
Silicate	0.312 ± 0.085	0.224 ± 0.075	0.464 ± 0.070	0.000 ± 0.000	-
Evaporite	0.432 ± 0.131	0.094 ± 0.040	0.474 ± 0.111	0.474 ± 0.111	-
Precipitation	0.036 ± 0.008	0.174 ± 0.007	0.790 ± 0.009	0.908 ± 0.041	-
Normalized to $\text{Ca}^{2+} + \text{Mg}^{2+} + \text{Na}^+ + \text{SO}_4^{2-}$ ($\Sigma^\pm$), results shown in Figures S1-S3					
Carbonate	0.660 ± 0.084	0.340 ± 0.084	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Silicate	0.312 ± 0.085	0.224 ± 0.075	0.464 ± 0.070	0.000 ± 0.000	0.000 ± 0.000
Evaporite	0.278 ± 0.072	0.063 ± 0.030	0.318 ± 0.104	0.318 ± 0.104	0.341 ± 0.052
Precipitation	0.033 ± 0.007	0.159 ± 0.006	0.721 ± 0.009	0.830 ± 0.038	0.087 ± 0.007
Sulfide ox.	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	1.000 ± 0.000

Normalizing to the sum of Ca^{2+} , Mg^{2+} , and Na^+ represents the simplest example of the differences in inversion results that arise from using the variance-targeting sampling algorithm instead of the n-1 sampling algorithm. However, excluding SO_4^{2-} from the inversion and neglecting to include an end-member for sulfide oxidation does mean that the outcomes of the example inversions are largely unrealistic. A more complete inversion, including an end-member representing sulfide oxidation and with SO_4^{2-} included in the normalization variable, is thus included in the supplementary materials (Figs. S1-S3; Table 1). The main difference between the two sets of inversion results is that changes in the evaporite end-member due to shifts in the sampling algorithm become less important when the inversion also includes SO_4^{2-} ; this result reflects that the inclusion of SO_4^{2-} itself generally limits the contribution of the evaporite end-



569 member relative to inversions without SO_4^{2-} included. We present the simpler case below in order
 570 to highlight the impact of the sampling algorithm on inversion results rather than to present the
 571 most accurate inversion for tabulating environmental fluxes.

572 The variance-targeting sampling algorithm produces large changes in the chemical priors
 573 for some end-members, such as the silicate end-member, and minimal changes in others, such as
 574 the carbonate end-member, relative to sampling with the n-1 algorithm (Fig. 5). For the silicate
 575 end-member under the n-1 routine, where Na^+ is used as the closure ion, the Na^+/Σ^+ range reaches
 576 both very high and low values (Fig. 5b, orange). In contrast, the variance-targeting algorithm
 577 generates a cloud of possible chemical compositions without the extreme Na^+/Σ^+ values (Fig. 5b,
 578 green). Whereas the standard deviation of the distribution of silicate Na^+/Σ^+ values is equal to 0.113
 579 for the n-1 algorithm, higher than the supplied value (Table 1), the standard deviation of silicate
 580 Na^+/Σ^+ values is 0.070 for the variance-targeting approach, in agreement with expectations (Table
 581 1). Similarly, the evaporite end-member no longer reaches very high Na^+/Σ^+ ratios under the
 582 variance-targeting algorithm and has reduced Mg^{2+}/Σ^+ ratios at high Ca^{2+}/Σ^+ values (Fig. 5c).
 583 While the evaporite end-member Na^+/Σ^+ distribution is expected to have a standard deviation of
 584 0.111 (Table 1), and the standard deviation of evaporite Na^+/Σ^+ is 0.110 under the variance-
 585 targeting sampling, the standard deviation of evaporite Na^+/Σ^+ is 0.136 under the n-1 sampling
 586 algorithm. The reduction in the chemical ranges of the silicate and evaporite end-members is
 587 potentially important for the inversion results as it means that the end-member chemistry input to
 588 the Monte Carlo simulations is more tightly constrained under the variance-targeting sampling
 589 algorithm than under the n-1 sampling algorithm.

590

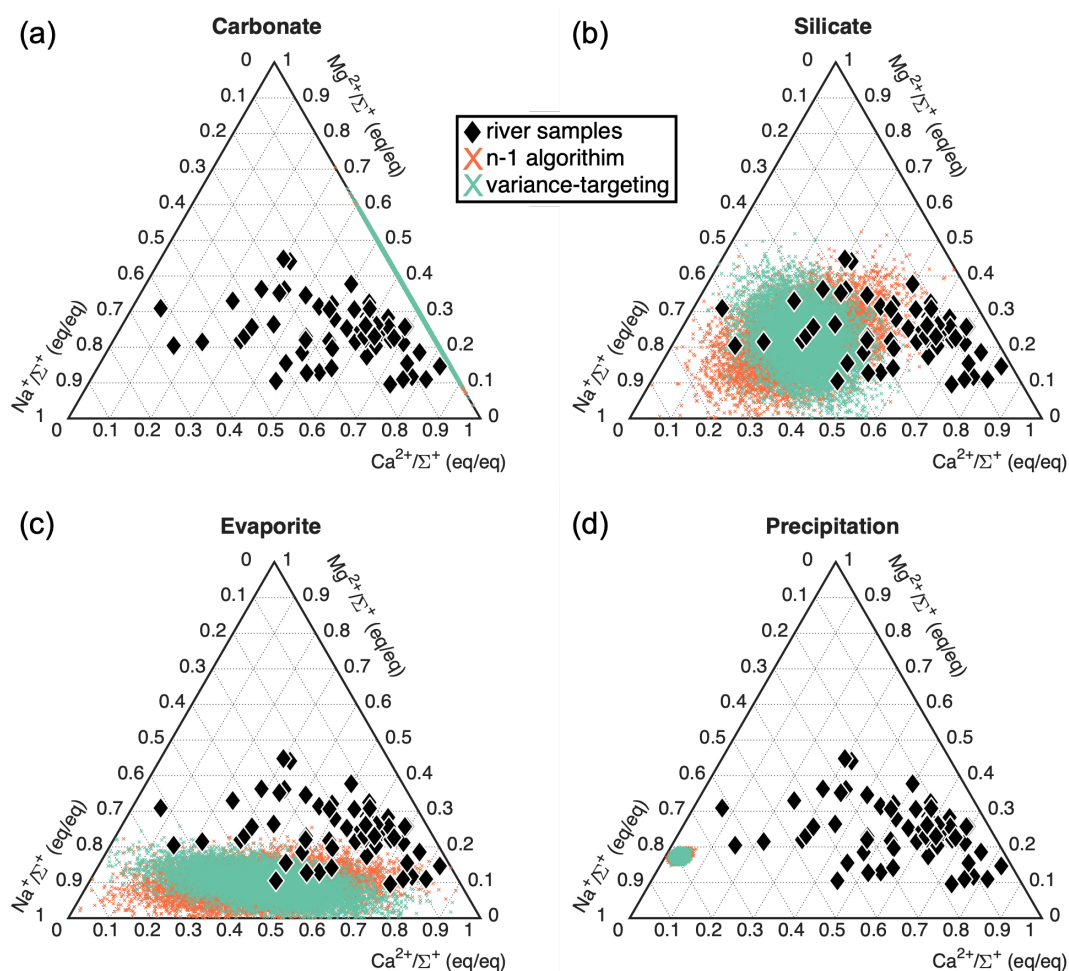
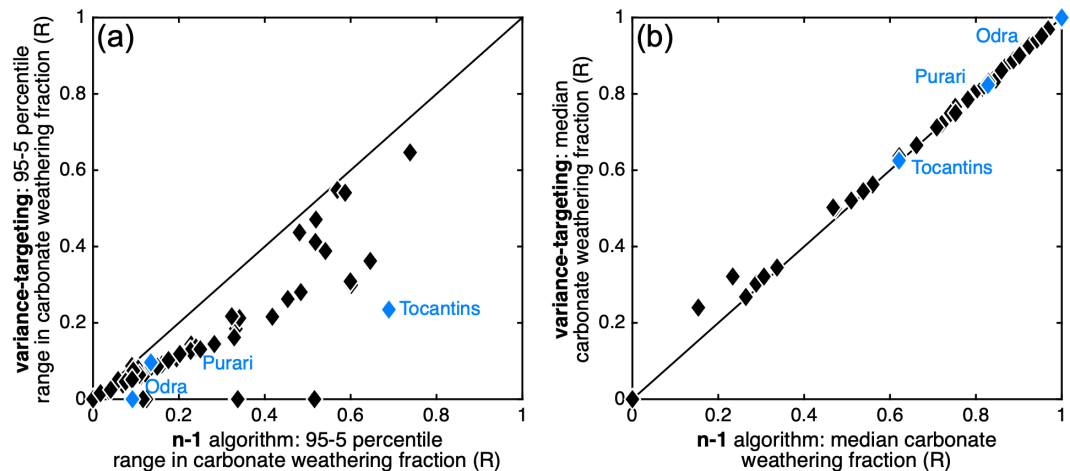


Fig. 5: Ca^{2+} - Mg^{2+} - Na^+ ternary diagrams with global river water samples (Gaillardet et al., 1999) and end-member chemical priors. The ranges of the (a) carbonate, (b) silicate, (c) evaporite, and (d) precipitation end-members are shown when calculated using the n-1 sampling algorithm (orange) or the variance-targeting algorithm (green). The difference between the two approaches is evident in the silicate end-member, where the reduction in Na^+/Σ^+ variance is manifest as a reduction in the density of high- Na^+/Σ^+ and low- Na^+/Σ^+ silicate end-member values. The chemistry of the evaporite end-member prior is also modified between the two algorithms, with a loss of high Na^+/Σ^+ values and a reduction in $\text{Mg}^{2+}/\Sigma^{2+}$ ratios at elevated $\text{Ca}^{2+}/\Sigma^{2+}$ ratios.



603 Inversion of the Gaillardet et al. (1999) data demonstrates that the variance-targeting
604 sampling algorithm results in a lower range for the fraction of carbonate mineral weathering,
605 commonly called R (Torres et al., 2016), than under the n-1 sampling routine (Fig. 6). Indeed, for
606 some sites, the 95th-5th percentile range of R is reduced from 0.33 (Elbe River), 0.34 (Seine River)
607 and 0.52 (Godavari River) to essentially zero. Moreover, the change in sampling algorithm has
608 only a small impact on the median of the results (Fig. 6b). These findings suggest that the change
609 in sampling algorithm modifies the tails of the distribution of the carbonate weathering fraction.
610 The dynamics reducing the range of R under variance-targeting end-member sampling are
611 demonstrated by evaluating changes in the distribution of R for the Tocantins River, Odra River,
612 and Purari River samples (Fig. 6, blue symbols).



613
614 **Fig. 6:** Comparison of inversion results for the carbonate weathering fraction (R) when sampling
615 end-member chemistry with either the n-1 algorithm or the variance-targeting algorithm. (a) The
616 95th-5th percentile range of the carbonate weathering fraction shows both modest and very large
617 changes between the sampling algorithms, with a reduced range of results under the variance-
618 targeting routine. (b) The median carbonate weathering fraction is very similar between the two
619 sampling strategies. Taken together, these results indicate that the variance-targeting sampling
620 algorithm removes the tails of the R distribution without major modifications to the median results.
621 Inversion results for the Tocantins River, the Odra River, and the Purari River (blue symbols) are
622 explored more thoroughly in Fig. 7.



623 We first consider changes in the distribution of R for the Tocantins River sample (Figs. 7a,
624 d, g). At this site, as well as in many others, the lowest values of R ($< 5^{\text{th}}$ percentile, triangular
625 plotting symbols) occur when the silicate end-member has a relatively low value for the Na^+/Σ^+
626 ratio and the highest values of R ($> 95^{\text{th}}$ percentile, square plotting symbols) occur when the silicate
627 end-member has a relatively high value for the Na^+/Σ^+ ratio (Fig. 7d). For both sampling
628 algorithms, high and low carbonate weathering fractions occur at a wide range of evaporite
629 chemistries (Fig. 7g). Because the variance-targeting sampling makes particularly high and low
630 silicate Na^+/Σ^+ ratios less probable (Fig. 5), the relative abundance of low-R and high-R results for
631 the Tocantins River sample decreases under the new sampling routine relative to the n-1 sampling
632 routine. In turn, the distribution of R values narrows (Fig. 7a) and both the $95^{\text{th}}\text{-}5^{\text{th}}$ percentile and
633 $75^{\text{th}}\text{-}25^{\text{th}}$ percentile ranges of R decrease. Overall, the extra variance on the silicate end-member
634 Na^+/Σ^+ ratio induced by the n-1 sampling algorithm forced the inversion results to have a wider
635 distribution than supported by the user-supplied end-member variances.

636 In the Odra River, a reduction in the $95^{\text{th}}\text{-}5^{\text{th}}$ percentile range of R is due to changes in the
637 distribution of the evaporite end-member chemistry. Unlike in the Tocantins River, high-R and
638 low-R simulations for the Odra River sample span the silicate compositional ternary diagram (Fig.
639 7e) while clustering in the evaporite ternary diagram (Fig. 7h). Under the n-1 sampling algorithm,
640 the lowest values of R occur at high values for evaporite Ca^{2+}/Σ^+ and Mg^{2+}/Σ^+ ratios (Fig. 7h). As
641 the evaporite chemistry giving rise to these low-R results becomes improbable under the variance-
642 targeting algorithm (Fig. 5c), the lowest values of R become larger, and the $95^{\text{th}}\text{-}5^{\text{th}}$ percentile range
643 of R is reduced to nearly zero as the distribution of R values converges to 1.

644 Results for the Purari River reflect changes in both the silicate end-member and evaporite
645 end-member (Figs. 7c, f, i). A subset of low-R results at particularly low evaporite Na^+/Σ^+ are only

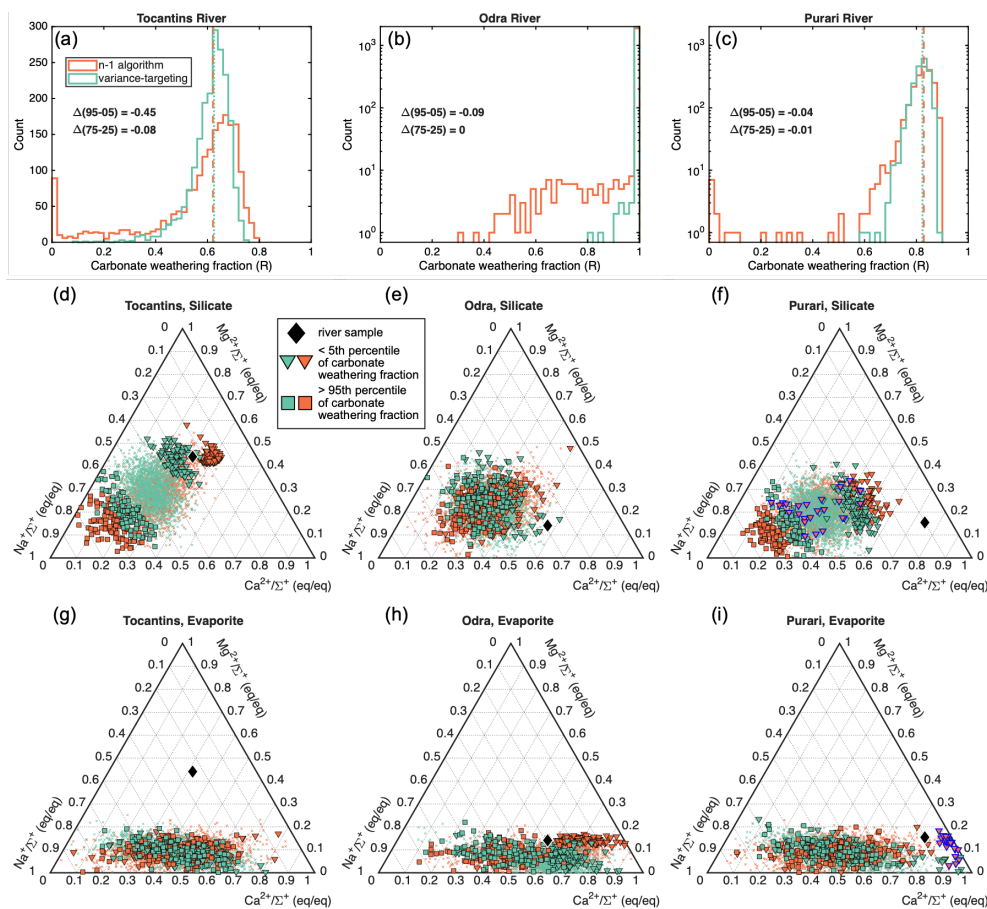


Fig. 7: Changes in the distribution of the carbonate weathering fraction (R) between the $n-1$ and variance-targeting sampling algorithms for water samples from the Tocantins River, Odra River, and Purari River (Gaillardet et al., 1999). The x, triangular, and square plotting symbols in the ternary diagrams all represent successful inversion results, where the triangles and squares highlight results below the 5th percentile and above the 95th percentile of R , respectively. Narrowing the range of end-member chemistry through improved sampling reduced the range of results. (a, b, c) Changes in the distribution of R for the three rivers across the two different sampling algorithms. The dotted lines indicate the median inversion result. (a, d, g) Changes in R for the Tocantins River reflect that low- R simulations tend to occur when the silicate end-member Na^+/Σ^+ ratio is low and high- R simulations tend to occur when the silicate end-member Na^+/Σ^+ ratio is high. Because the variance-targeting algorithm produces fewer instances of extreme silicate Na^+/Σ^+ ratios relative to the $n-1$ routine, the change in sampling reduces the 95th-5th percentile range of R . (b, e, h) Within the Odra River results, the loss of low- R inversion outcomes under the variance-targeting sampling algorithm reflects the removal of simulations with high Ca^{2+}/Σ^+ and high Mg^{2+}/Σ^+ evaporite chemistry. (c, f, i) Changes in the distribution of R for the Purari River reflect changes in both the silicate and evaporite end-members. Triangular symbols with blue edges in (f) and (i) highlight inversion results for the $n-1$ algorithm with both low R values and particularly low evaporite Na^+/Σ^+ ratios.



possible with the n-1 sampling routine (Fig. 7i, orange triangles with blue edges), which plot in the middle of the silicate compositional ternary (Fig. 7f), while the reduction in the silicate end-member Na^+/Σ^+ range also narrows the distribution of R through removal of high-R and low-R results. As alluded to above, evaporite-related differences in results for the Odra River and Purari River samples are largely removed when the inversion includes SO_4^{2-} (Fig. S3). In the more realistic inversion scenarios, reductions in the 95th-5th percentile range of R due to the change in sampling algorithm (Fig. S2) mostly reflect reductions in the chemical range of the silicate end-member.

5. Conclusions

River inversion models are essential for quantitatively allocating fluvial dissolved load among contributions from lithologic and process end-members. However, the standard approach in MEANDIR to sum-constrained sampling of Gaussian distributions artificially modifies the variance in a single chemical parameter for each end-member, typically to higher values. Here we introduce a new variance-targeting and sum-constrained sampling algorithm for normal chemical distributions. The updated algorithm seeks to maximize the entropy of the vector of end-member chemical distributions by maximizing the absolute value of a matrix determinant while subject to optimization constraints that account for target variances. The optimized matrix then transforms a vector of independent pulls from the standard normal distribution. While the approach preserves user-specified variances during the sampling, these variances are then modified when enforcing physical constraints. The new sampling routine is implemented in MEANDIR and used to invert the global river dataset of Gaillardet et al. (1999). The change in sampling algorithm reduces the likelihood of extreme end-member chemical values and, in turn, reduces the 95th-5th percentile



688 range of results for the fraction of carbonate mineral weathering. The median inversion result, in
689 contrast, does not change substantially under the new sampling algorithm. A major limitation of
690 the new sampling approach is that it does not address the case of user-supplied uniform
691 distributions, where the end-member chemical ratio used for closure will again have variance offset
692 from the user-supplied values under the n-1 sampling strategy. As a result, we recommend defining
693 the chemistry of sum-normalized lithologic or process end-members with Gaussian distributions
694 rather than uniform distributions within the MEANDIR river inversion model.

695

696 **6. Code and data availability**

697 The version of MEANDIR (v1.4) used to produce the results in this manuscript is archived
698 on Zenodo (Kemeny, 2026; DOI: 10.5281/zenodo.22257975) under a Creative Commons
699 Attribution license. No new data were generated for this work; data from Gaillardet et al. (1999)
700 are available through the MEANDIR download and the original publication.

701

702 **7. Author contributions**

703 PCK conceptualized the study, implemented the updated sampling algorithm in
704 MEANDIR, drafted the figures, and wrote the manuscript. OG developed the algorithm for
705 variance-targeting sum-constrained Gaussian sampling.

706

707 **8. Competing interests**

708 The authors declare that they have no conflict of interest.

709

710



711 9. Acknowledgments

712 This article emerged from a discussion at the 2025 Hertz Summer Workshop and both
713 authors are grateful to the Fannie and John Hertz Foundation for organizing the symposium. PCK
714 was partially supported during this research through the Investment in Science Fund at the Woods
715 Hole Oceanographic Institution. ChatGPT (5.2 Thinking) was utilized to improve the
716 implementation of the variance-targeting sum-constrained sampling algorithm within MEANDIR.
717 Specifically, engagement with ChatGPT led to the conclusion that employing the sum of the
718 logarithms of the singular values of $\mathbf{G}$ would generate a more stable cost function than maximizing
719 $|\det(\mathbf{G})|$ directly. Second, ChatGPT helped us diagnose that numerical instability during early
720 implementation of the algorithm derived from differences in variable scaling, prompting us to split
721 the matrix $\mathbf{G}$ into the product of $\mathbf{D}$ and $\mathbf{X}$. We confirmed each of these changes improved the
722 numerical robustness of the code.



10. References

- Altmann, Y., McLaughlin, S. and Dobigeon, N., 2014, June. Sampling from a multivariate Gaussian distribution truncated on a simplex: a review. In 2014 IEEE Workshop on Statistical Signal Processing (SSP) (pp. 113-116). IEEE.
- Berner, R.A. and Caldeira, K., 1997. The need for mass balance and feedback in the geochemical carbon cycle. *Geology*, 25(10), pp.955-956.
- Berner, R.A., Lasaga, A.C. and Garrels, R.M., 1983. Carbonate-silicate geochemical cycle and its effect on atmospheric carbon dioxide over the past 100 million years. *Am. J. Sci. (United States)*, 283(7).
- Burke, A., Present, T.M., Paris, G., Rae, E.C., Sandilands, B.H., Gaillardet, J., Peucker-Ehrenbrink, B., Fischer, W.W., McClelland, J.W., Spencer, R.G. and Voss, B.M., 2018. Sulfur isotopes in rivers: Insights into global weathering budgets, pyrite oxidation, and the modern sulfur cycle. *Earth and Planetary Science Letters*, 496, pp.168-177.
- Calmels, D., Gaillardet, J., Brenot, A. and France-Lanord, C., 2007. Sustained sulfide oxidation by physical erosion processes in the Mackenzie River basin: Climatic perspectives. *Geology*, 35 (11), pp.1003-1006.
- Gaillardet, J.D., Dupré, B., Louvat, P. and Allegre, C.J., 1999. Global silicate weathering and CO₂ consumption rates deduced from the chemistry of large rivers. *Chemical geology*, 159(1-4), pp.3-30.
- Hemingway, J.D., Olson, H., Turchyn, A.V., Tipper, E.T., Bickle, M.J. and Johnston, D.T., 2020. Triple oxygen isotope insight into terrestrial pyrite oxidation. *Proceedings of the National Academy of Sciences*, 117(14), pp.7650-7657.
- Hilton, R.G. and West, A.J., 2020. Mountains, erosion and the carbon cycle. *Nature Reviews Earth & Environment*, 1(6), pp.284-299.
- Kemeny, P.C. and Torres, M.A., 2021. Presentation and applications of mixing elements and dissolved isotopes in rivers (MEANDIR), a customizable MATLAB model for Monte Carlo inversion of dissolved river chemistry. *American Journal of Science*, 321(5), pp.579-642.
- Kemeny, P.C., Torres, M.A., Fischer, W.W. and Blättler, C.L., 2024. Balance and imbalance in biogeochemical cycles reflect the operation of closed, exchange, and open sets. *Proceedings of the National Academy of Sciences*, 121(12), p.e2316535121.
- Kemeny, P.C. (2026) "MEANDIR v1.4". Zenodo. <https://doi.org/10.5281/zenodo.22257975>.
- Moon, S., Chamberlain, C.P. and Hilley, G.E., 2014. New estimates of silicate weathering rates and their uncertainties in global rivers. *Geochimica et Cosmochimica Acta*, 134, pp.257-274.
- Négrel, P., Allegre, C.J., Dupré, B. and Lewin, E., 1993. Erosion sources determined by inversion of major and trace element ratios and strontium isotopic ratios in river water: the Congo Basin case. *Earth and Planetary Science Letters*, 120(1-2), pp.59-76.
- Spence, J. and Telmer, K., 2005. The role of sulfur in chemical weathering and atmospheric CO₂ fluxes: evidence from major ions, $\delta^{13}\text{C}_{\text{DIC}}$, and $\delta^{34}\text{S}_{\text{SO}_4}$ in rivers of the Canadian Cordillera. *Geochimica et Cosmochimica Acta*, 69(23), pp.5441-5458.
- Stallard, R.F. and Edmond, J.M., 1983. Geochemistry of the Amazon: 2. The influence of geology and weathering environment on the dissolved load. *Journal of Geophysical Research: Oceans*, 88(C14), pp.9671-9688.
- Torres, M.A., West, A.J., Clark, K.E., Paris, G., Bouchez, J., Ponton, C., Feakins, S.J., Galy, V. and Adkins, J.F., 2016. The acid and alkalinity budgets of weathering in the Andes–



769 Amazon system: Insights into the erosional control of global biogeochemical cycles. Earth
770 and Planetary Science Letters, 450, pp.381-391.
771 West, A.J., Galy, A. and Bickle, M., 2005. Tectonic and climatic controls on silicate weathering.
772 Earth and Planetary Science Letters, 235(1-2), pp.211-228.
773 Zondervan, J.R., Hilton, R.G., Dellinger, M., Clubb, F.J., Roylands, T. and Ogrič, M., 2023. Rock
774 organic carbon oxidation CO₂ release offsets silicate weathering sink. Nature, 623(7986),
775 pp.329-333.