Geophysical Research Letters

RESEARCH LETTER
10.1029/2023GL104147

Key Points:

- Tracer fluctuations contain information that can help constrain mixing models, including those with overlapping end-member signatures
- This approach can also constrain mixing models with many sources but only one tracer, or those in which some sources may be missing
- The concepts and math behind this approach are outlined, and validated using benchmark tests; a script for the method is provided

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. W. Kirchner,
kirchner@ethz.ch

Citation:

Kirchner, J. W. (2023). Mixing models with multiple, overlapping, or incomplete end-members, quantified using time series of a single tracer. *Geophysical Research Letters*, *50*, e2023GL104147. https://doi.org/10.1029/2023GL104147

Received 15 APR 2023
Accepted 18 JUN 2023

Author Contributions:

Conceptualization: James W. Kirchner
Formal analysis: James W. Kirchner
Methodology: James W. Kirchner
Software: James W. Kirchner
Validation: James W. Kirchner
Writing – original draft: James W. Kirchner
Writing – review & editing: James W. Kirchner

Mixing Models With Multiple, Overlapping, or Incomplete End-Members, Quantified Using Time Series of a Single Tracer

James W. Kirchner[1,2,3]

[1]Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland, [2]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland, [3]Department of Earth and Planetary Science, University of California, Berkeley, CA, USA

Abstract Mixing models are used throughout earth and environmental science to quantify the relative contributions of sources to mixtures, based on chemical or isotopic tracers. Often, however, some end-members are missing or their tracer distributions overlap, precluding the use of conventional mixing models. Here I show how these constraints can be overcome by exploiting the information contained in tracer time-series fluctuations. This approach, ensemble end-member mixing analysis (EEMMA), can potentially quantify many sources using a single tracer, even if their mean concentrations are indistinguishable. EEMMA can also quantify source contributions when some sources are unknown, and even infer the tracer time series of a missing source. Benchmark tests with synthetic data verify the reliability of this approach, thus expanding the range of mixing models that can be quantified using tracer time series. An R script is provided for the necessary calculations, including error propagation.

Plain Language Summary Chemical and isotopic tracers are widely used to partition mixtures (such as air masses, stream flows, xylem water, or biogeochemical fluxes) among potential contributing sources. These "mixing models" are widely used to improve process understanding throughout Earth and environmental science. However, because conventional mixing models are based on comparing the mean tracer concentrations in the sources and mixture, they typically require that all sources have been identified and sampled, their average tracer concentrations are distinctly different, and the number of sources does not exceed the number of independent tracers, plus one. Temporal fluctuations in tracer concentrations can provide additional information that loosens these constraints, thus expanding the range of feasible mixing models. Here I outline this approach and test its reliability using synthetic benchmark data. With this approach, in contrast to conventional mixing models, a single tracer can be used to quantify contributions from many different sources, even if their means overlap. Even when some sources are unmeasured or even unknown, this approach can estimate the remaining sources' mixing fractions, and even, given sufficiently accurate data, can estimate the mixing fraction and tracer signature of an unknown source. Thus, this approach expands the range of applications for mixing models.

1. Introduction: Conventional Versus Ensemble End-Member Mixing Analysis

Mixing models are widely used for flux estimation and source attribution in fields ranging from marine and atmospheric science to hydrology, ecophysiology, and biogeochemistry (e.g., Gessler et al., 2014; Jiskra et al., 2021; Klaus & McDonnell, 2013; Layman et al., 2012; Tetzlaff et al., 2014). Mixing models infer the relative contributions of sources (also termed "end-members") to mixtures, using conservative tracers such as passive solutes or stable isotopes. However, conventional mixing models (e.g., Beria et al., 2020; Christophersen et al., 1990; Stock et al., 2018) require that all of the sources are known and sampled, and that their mean tracer signatures do not overlap. They further require that the number of end-members is less than the number of tracers, plus one (and potentially much less, when tracers are correlated; Barthold et al., 2011; Christophersen et al., 1990). The purpose of this paper is to show how all of these constraints can be overcome by exploiting the information contained in tracer fluctuations.

The simplest mixing model concerns a single tracer in two sources and a mixture, such as catchment storage and recent precipitation as sources contributing to streamflow, or soil water and groundwater as sources contributing to plant water uptake. If the tracer concentrations in the mixture and the two sources are m, v_1, and v_2, respectively, then mass conservation implies that

$$m = f_1 v_1 + f_2 v_2, \tag{1}$$

where f_1 and f_2 are the fractions of the mixture that are derived from sources 1 and 2. The mass balance of Equation 1 inherently assumes that the two sources jointly account for all of the mixture, and thus that $f_1 + f_2 = 1$. Thus Equation 1 can be rewritten as

$$m = f_1 v_1 + (1 - f_1)\, v_2, \tag{2}$$

leading to the conventional end-member mixing equation

$$\hat{f}_1 = \frac{m - v_2}{v_1 - v_2}, \quad \hat{f}_2 = 1 - \hat{f}_1, \tag{3}$$

which can be applied either to individual measurements of m, v_1, and v_2, or to their averages. From Equation 3, one can see that the estimated mixing fractions $\hat{f}_1$ and $\hat{f}_2$ will be highly uncertain unless the two end-members v_1 and v_2 are much farther apart than their own uncertainties. One can also see that because the mixture measurements m are also uncertain, they can lie outside the interval between v_1 and v_2, leading to physically impossible values (less than zero or greater than one) for the mixing fractions f_1 and f_2.

Figure 1 illustrates this issue, using three scenarios shown in three rows of panels. The three scenarios are identical in every respect (e.g., the mixing fraction is the same and the tracer fluctuations in both the sources and the mixture are the same), except the means of the sources are far apart in the top row and close together in the bottom row. As the means of the two sources converge, the mixing fractions estimated by conventional mixing models (Equation 3) become highly uncertain (Figures 1b, 1e, and 1h), due to fluctuations in the sources and the mixture.

By contrast, ensemble end-member mixing analysis (EEMMA) exploits these fluctuations and their correlations to estimate mixing fractions, based on the observation that if multiple sets of tracer measurements are available, the mixing model of Equations 1–3 can be recast as a no-intercept regression equation,

$$(m_i - v_{i,2}) = f_1\,(v_{i,1} - v_{i,2}), \tag{4}$$

or equivalently

$$y_i = \beta\, x_i + e_i, \tag{5}$$

where $y_i = m_i - v_{i,2}$, $x_i = v_{i,1} - v_{i,2}$, the subscript i denotes individual time steps, the regression slope β estimates f_1, and the residual term e_i subsumes any error in y_i. This regression-based approach requires multiple sets of measurements of m, v_1, and v_2, which are often available in tracer studies. Equations 4 and 5 also assume that f_1 is a constant, although as shown in Section 3 below, it will also estimate the average of f_1 if the mixing fractions are time-varying.

EEMMA has the distinct advantage that points with $v_1 \approx v_2$, which blow up Equation 3, have virtually no leverage on the solution to Equations 4 and 5. Thus the mixing fraction f_1 can be reliably estimated even as the average difference between v_1 and v_2 shrinks to zero (Figures 1c, 1f, and 1i). When the means of v_1 and v_2 are much farther apart than their fluctuations (top row of Figure 1), the mean displacements between m, v_1, and v_2 reliably estimate the mixing fraction f_1 in both conventional mixing models (Figure 1b) and EEMMA (Figure 1c). But as these mean displacements become small relative to the fluctuations, conventional mixing models become much less reliable than EEMMA (compare Figures 1h and 1i), because Equations 4 and 5 can exploit the information contained in both the mean displacements and the fluctuation correlations between the sources and the mixture. Mathematically this can be seen by decomposing each of the tracer concentrations into their means and fluctuations around those means, recasting Equation 4 as:

$$\left(\overline{m} - \overline{v}_2\right) + \left(m'_i - v'_{i,2}\right) = f_1 \left[\left(\overline{v}_1 - \overline{v}_2\right) + \left(v'_{i,1} - v'_{i,2}\right)\right], \tag{6}$$

where overbars and primes indicate means and fluctuations, respectively. From Equation 6, one can see that when the differences in means are large compared to the differences in fluctuations, they will have most of the influence on the solution (e.g., Figures 1a–1c), whereas the converse will be true when the differences in fluctuations are large compared to the differences in means (e.g., Figures 1g–1i).

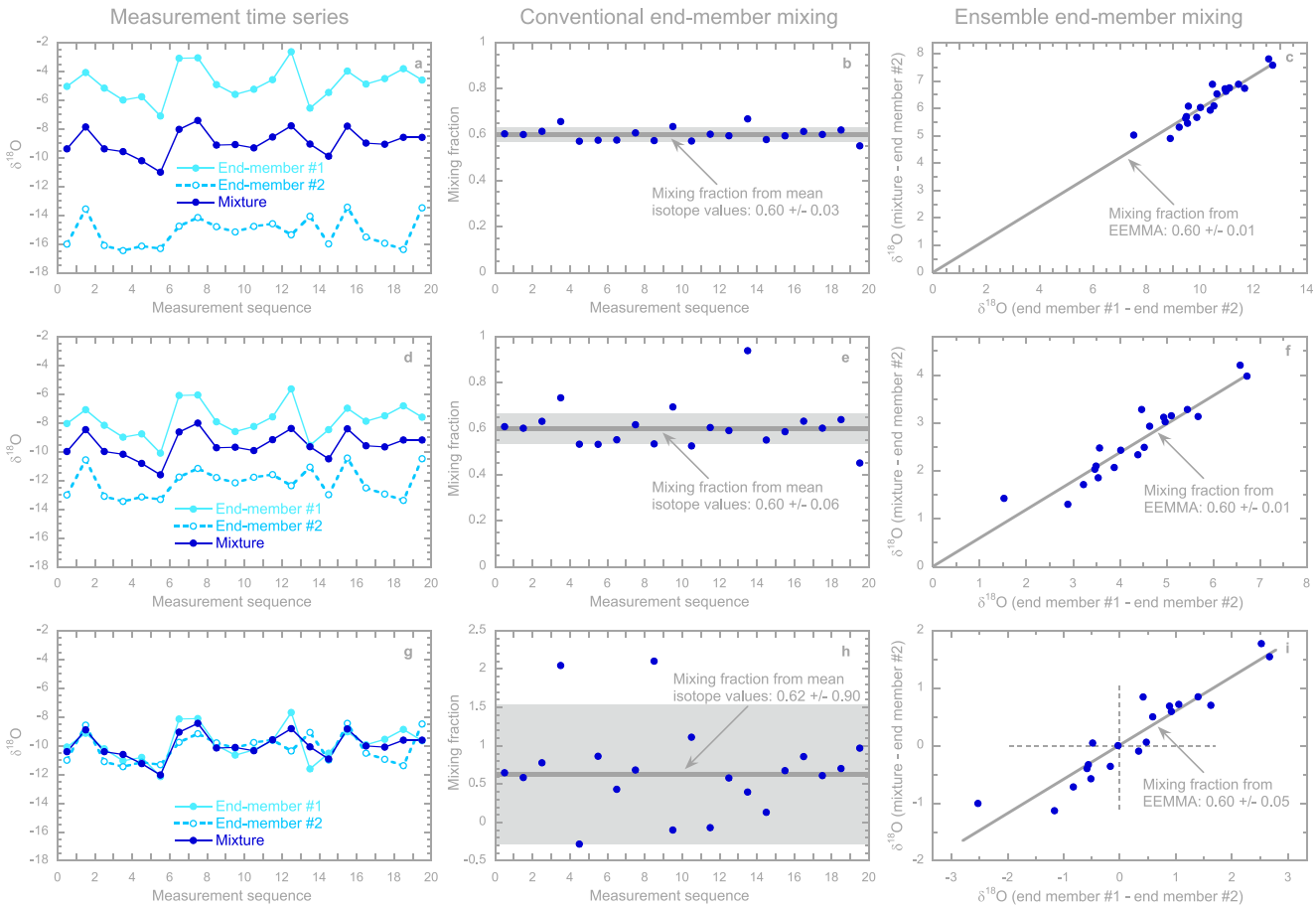

Figure 1. Conventional and ensemble end-member mixing, illustrated using three contrasting scenarios of synthetic stable isotope data. All three scenarios (one per row of panels) have identical mixing ratios (0.6 and 0.4 for end-members 1 and 2, respectively) and identical tracer fluctuations, including identical measurement errors (RMSE = 0.25‰) in the mixture. Only the offsets between the end-member means are different, shrinking from ~10‰ in the top row to ~4‰ in the middle row and 0.3‰ in the bottom row. The middle panels (b, e, h) show mixing fractions obtained by conventional end-member mixing (Equation 3) applied to each measurement individually (blue points) and to the means of each measurement series (dark gray line and light gray band indicating the estimated mixing fraction and its standard error). The right-hand panels illustrate ensemble end-member mixing (Equation 4), which estimates average mixing fractions from zero-intercept regression slopes. As the end-member means converge (i.e., from top to bottom), conventional end-member mixing becomes unreliable, even exceeding the physically possible range of 0–1 (h; note the large uncertainty range in gray). By contrast, ensemble end-member mixing (c, f, i) uses the fluctuation correlations between the mixture and end-members as additional constraints, reliably estimating the mixing fractions of isotope time series whose means overlap.

2. Quantifying Multiple Mixing Fractions Using a Single Tracer

Equation 4 can be straightforwardly generalized to a zero-intercept multiple regression that quantifies the mixing fractions of multiple sources using a single tracer:

$$(m_i - v_{i,k}) = f_1(v_{i,1} - v_{i,k}) + \cdots + f_{k-1}(v_{i,k-1} - v_{i,k}). \quad (7)$$

In principle the number of sources k is limited only by the available degrees of freedom in the measurements. At most two of these sources can be constants, since with just one tracer, differences between end-member means can constrain at most one mixing fraction. But the same limitation does not apply to correlations; here, instead, the theoretical limit depends primarily on the number of measurements rather than the number of tracers. The practical limit in real-world cases (as revealed by the standard errors of the mixing fractions) will also depend on how correlated the end-members are and how noisy the measurements are (as is also true of conventional mixing models; Barthold et al., 2011). Figure 2 shows results of 1,000 benchmark tests using 20 measurements of a single tracer to estimate the mixing fractions of 2, 4, and 6 sources. In each test the means of the end-members exactly overlap; this makes the estimation task harder because it forces Equation 7 to rely solely on the fluctuations. Nevertheless, as Figure 2 shows, the approach outlined here can reliably estimate multiple mixing fractions, and their standard errors, using a single tracer.

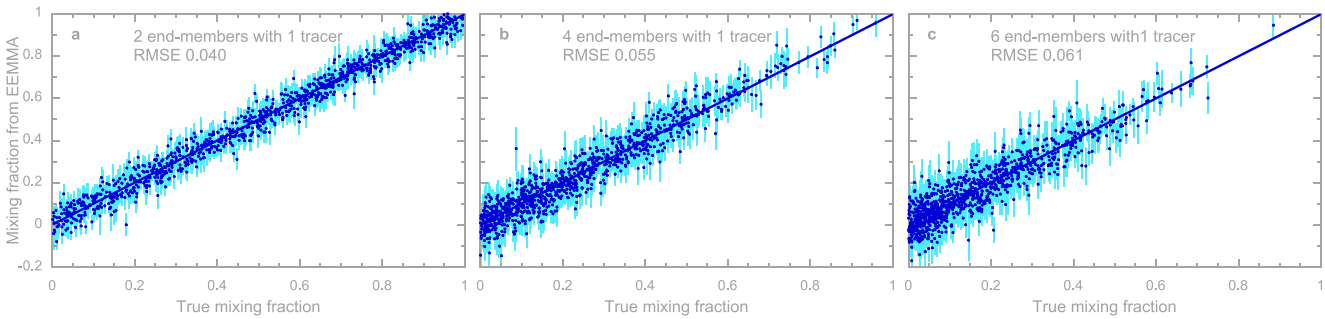

Figure 2. Benchmark tests of ensemble end-member mixing, using a single tracer to quantify mixing fractions of 2, 4, and 6 end-members (a, b, and c, respectively). Each plot has 1,000 points; each point shows the mixing fraction of the first end-member, obtained from analyzing a random synthetic data set of 20 time steps (e.g., Figure 1). Error bars (light blue) indicate one standard error. True mixing fractions are randomly chosen, so values near 1 become rarer as the number of end-members increases. All end-members are uncorrelated, with means = 0 and standard deviations = 1. The mixture contains random noise with standard deviation = 0.25, so the signal-to-noise ratios in the mixture decrease from 8 to 4 and 2.7 in panels a, b, and c, respectively (because the signal variance in the mixture decreases proportionally to the number of end-members). Estimated mixing fractions from Equations 4 and 7 generally conform closely to the true mixing fractions, with standard errors typically intersecting the 1:1 line. Thus both the mixing fractions and their uncertainties are accurately quantified.

3. Time-Varying Mixing Fractions

Tracer time series are often interpreted as reflecting time-varying mixing fractions, with end-member mixing being used to estimate mixing fractions for each time step. As the blue points in Figures 1b, 1e, and 1h show, these inferred mixing fractions can become unstable unless there is clear separation between the end-members. Ensemble end-member mixing can reliably estimate a single, ensemble estimate of the mixing fraction in such cases (Figure 1i), but its relationship to the individual time-varying mixing fractions is not intuitively obvious. However, one can show that, barring strong linear or nonlinear correlations between fluctuations in these mixing fractions and the end-member tracer concentrations themselves, their average will be closely approximated by the ensemble mixing fraction, as estimated by Equations 4 and 7 (see Supporting Information S1).

Figure 3 shows scenarios similar to those in Figure 1, but with mixing fractions that vary through time. As Figure 3 illustrates, time-varying mixing fractions (orange lines, Figures 3b, 3e, and 3h) make tracers in the mixture more variable (dark blue lines, Figures 3a, 3d, and 3g), and thus add scatter to the EEMMA regressions (Figures 3c, 3f, and 3i). Nonetheless EEMMA closely approximates the true average mixing fraction of 0.6 for source #1, even when the end-members overlap and conventional end-member mixing becomes unreliable (compare the uncertainties in Figures 3h and 3i). Benchmark tests show that EEMMA approximates the average mixing fraction in time-varying systems, including those with more than two sources (Figure S1 in Supporting Information S1). Although EEMMA cannot estimate mixing fractions for individual time steps, it can be applied to contrasting time periods (e.g., comparing different seasons or weather patterns), thus quantifying how mixing fractions differ among contrasting ambient conditions.

4. Incomplete Mixing Models and Unmeasured End-Members

Conventional end-member mixing analysis requires measurements for all the end-members (meaning all sources that contribute to the mixture in the real world, not just in a theory, a model, or a sampling program). This requirement often cannot be met.

By exploiting information contained in correlations between sources and mixtures, by contrast, EEMMA can estimate mixing fractions even when some end-members are missing. Consider a case with tracer measurements in a mixture m and one or more sources $v_1 \ldots v_k$. The mixture may also reflect unmeasured sources (either time-varying or constant). These unmeasured sources can be combined into a single end-member u, representing their volume-weighted average.

Mass conservation implies that at each time step i ($i = 1 \ldots n$),

$$m_i = f_1 v_{i,1} + \cdots + f_k v_{i,k} + (1 - f_1 \cdots - f_k)\, u_i, \tag{8}$$

where $f_1 \ldots f_k$ are the fractions of the mixture that are derived from the measured end-members $v_1 \ldots v_k$, and $(1 - f_1 \ldots - f_k)$ is the mixing fraction f_u of the unmeasured end-member u. If the unmeasured end-member u_i is

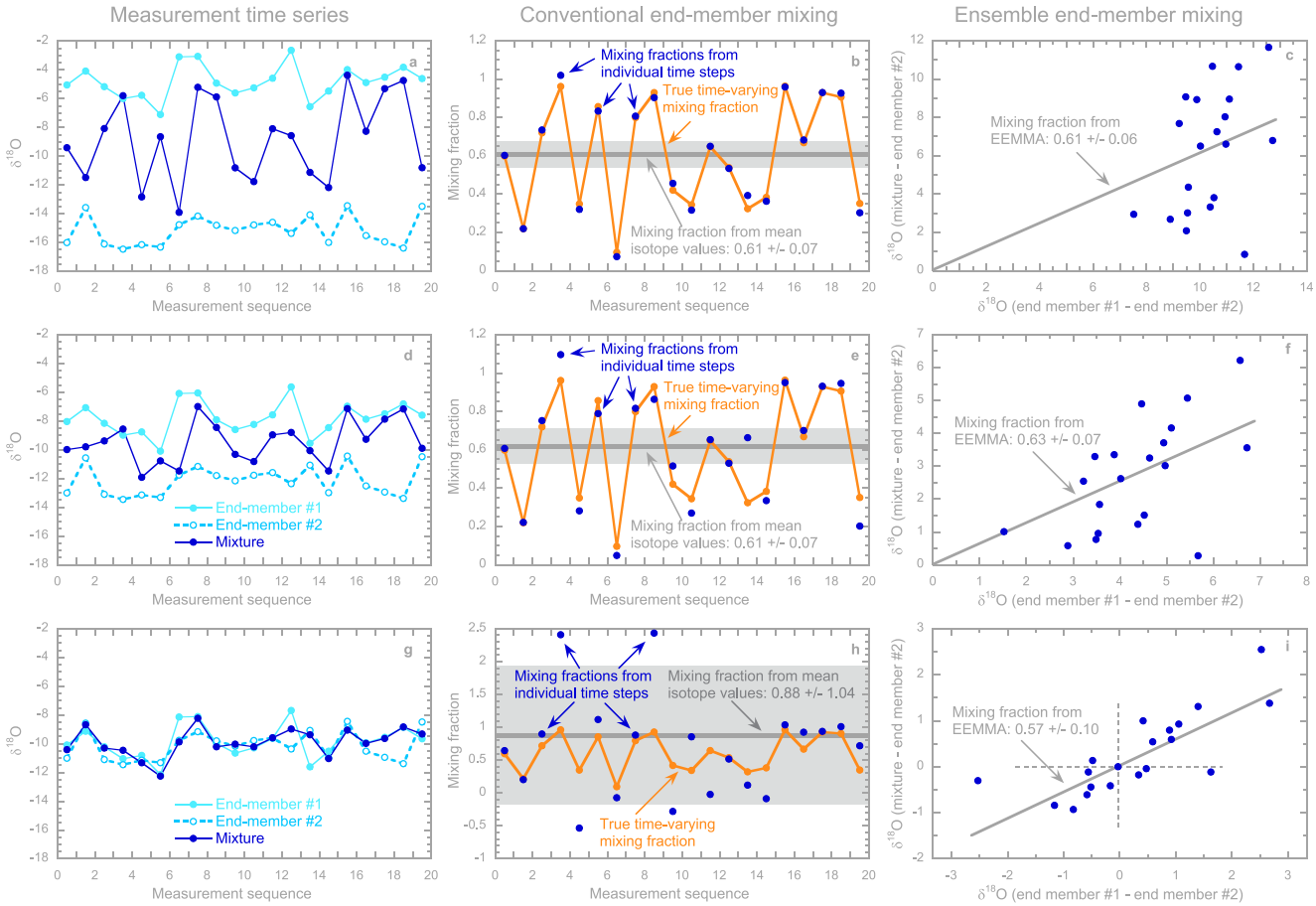

Figure 3. Conventional and ensemble end-member mixing when mixing fractions vary through time. Three scenarios are shown in the three rows of panels. These are identical to the scenarios in Figure 1, except the mixing fraction of source #1 varies between 0.10 and 0.96, with an average of 0.6 (orange lines; these are identical in all three rows, but look different because the axes differ). Only the offsets between the end-member means vary, shrinking from ~10 to ~4‰ and 0.3‰ in the top, middle, and bottom rows, respectively. The middle panels (b, e, h) show mixing fractions obtained by conventional end-member mixing (Equation 3) applied to each time step individually (blue points) and to the means of each measurement series (dark gray line and light gray band indicating the estimated mixing fraction and its standard error). The right panels illustrate ensemble end-member mixing (Equation 4), which estimates average mixing fractions from zero-intercept regression slopes. Conventional end-member mixing accurately identifies the mixing fractions of individual time steps when the end-member means are far apart (b), but becomes unreliable as the means converge (h; note the large uncertainty range indicated in gray). The zero-intercept regressions of ensemble end-member mixing (c, f, i) cannot estimate mixing fractions for individual time steps, but reliably estimate their averages, even when the end-members substantially overlap.

uncorrelated with the measured end-members v_{ij}, Equation 8 is equivalent to the conventional linear regression equation

$$m_i = \sum_{j=1}^{k} v_{i,j}\beta_j + \alpha + e_i, \tag{9}$$

where each regression slope β_j estimates the corresponding f_j, the regression intercept α estimates $(1 - \Sigma f_j)\,(\bar{u})$, and the residual term e_i subsumes $(1 - \Sigma f_j)\,u'_i$ plus any error in m_i. (In contrast to Equations 4–7, Equation 9 requires an intercept, to account for the unmeasured end-member.) One can also estimate the time series of the missing end-member as

$$\hat{u}_i = \frac{\alpha + e_i}{1 - \sum_{j=1}^{k} f_j}. \tag{10}$$

The reliability of such estimates will depend primarily on how much of the residual is due to the effects of the unknown end-member versus random sampling and measurement error. As Figures 4a and 4b and Figures S2a

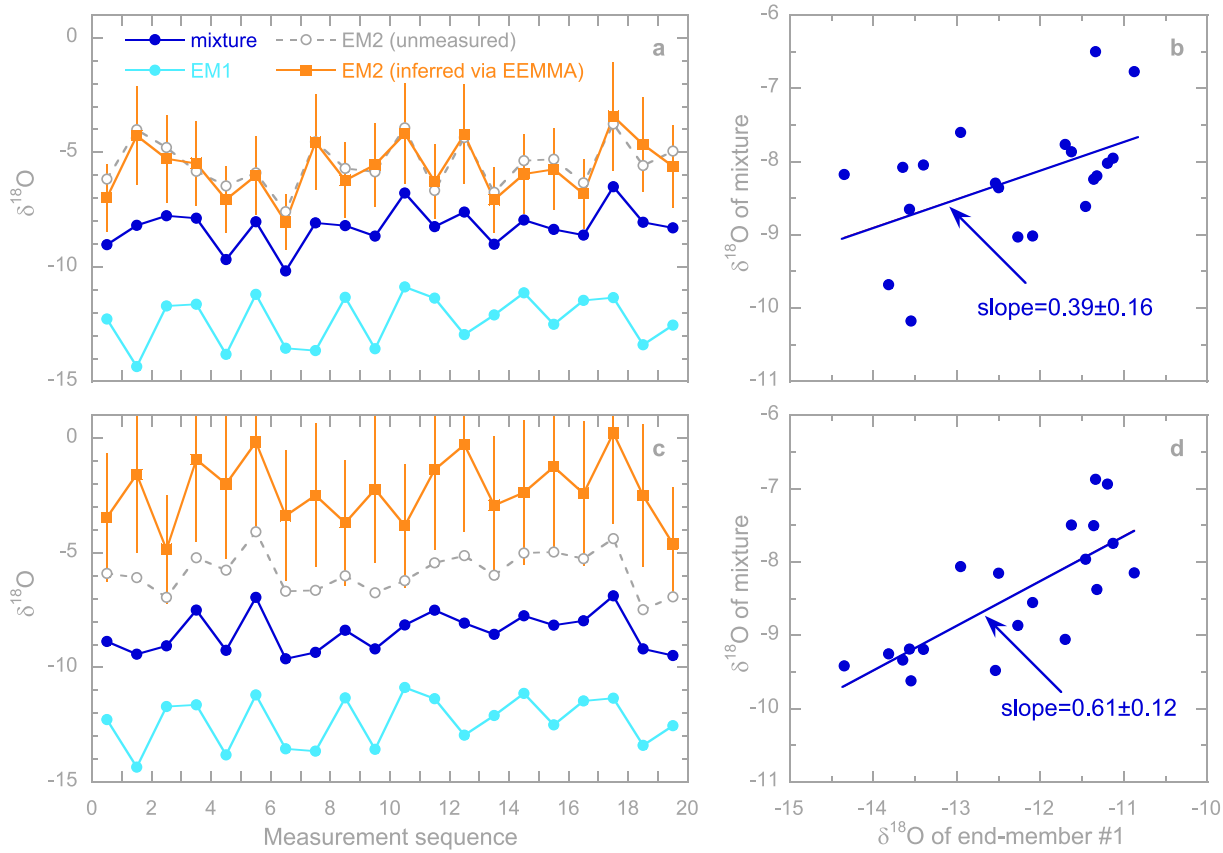

Figure 4. Estimation of mixing fractions when an end-member is unmeasured. In (a) and (c), end-member #1 (light blue) contributes 40% to the mixture (dark blue), and end-member #2 (gray circles), which is unmeasured, contributes 60%. When the two end-members are uncorrelated ($r < 0.05$: top panels), ensemble end-member mixing analysis correctly estimates the mixing fraction of the measured end-member (b), and correctly infers the unmeasured end-member time series (orange squares, panel a) via Equation 10. However, when the two end-members are strongly correlated ($r \approx 0.5$: bottom panels), the unmeasured end-member inflates the correlation and regression slope of the measured end-member with the mixture (d), and thus inflates its apparent mixing fraction. Thus the mixing fraction of the unmeasured end-member is underestimated, and its inferred time series (orange squares, panel c) lies far from the true values (gray circles), although they are strongly correlated ($r \approx 0.75$).

and S2b in Supporting Information S1 show, this approach can accurately infer mixing fractions for measured and unmeasured sources, and accurately reconstruct an unmeasured source's time series.

However, the assumption that the unmeasured end-members are not strongly correlated with the measured ones is crucial, because such correlations can distort the estimated mixing fractions of the measured sources. Consider the simple case of two potentially correlated end-members, one (v_1) that has been measured and another (v_2) that has not. Propagation of covariance leads directly to the following formula for the expected regression slope when the mixture m is plotted as a function of the measured end member v_1 alone:

$$\mathrm{E}\left(\beta_{m,v_1}\right) = f_1 + f_2\, r_{v_1,v_2}\, \frac{s_{v_2}}{s_{v_1}}, \tag{11}$$

where s_{v_1} and s_{v_2} are the standard deviations of v_1 and v_2, and r_{v_1,v_2} is the correlation between them. Figures 4c and 4d illustrates how positive correlations between measured and unmeasured sources lead to exaggerated estimates of the mixing fraction of the measured source, and thus to distorted inferences concerning the unmeasured source. Figure S2 in Supporting Information S1 shows that Equations 8 and 9 accurately estimate mixing fractions for measured and unmeasured sources that are uncorrelated, and that Equation 11 correctly describes how these mixing fractions are distorted by correlations between measured and unmeasured sources.

It will normally be impossible to empirically verify whether the measured end-members are correlated with any unmeasured ones, precisely because they are unmeasured. Testing for correlations between $\hat{u}_i$ (Equation 10) and

the measured end-members won't help; these correlations are always zero, whether or not the true unmeasured end-member is similarly uncorrelated. Nevertheless, in assuming that any unmeasured end-members are uncorrelated with the measured ones, EEMMA is less restrictive than conventional mixing models, which require that no unmeasured end-members exist at all. If Σf_j is statistically indistinguishable from 1, the set of end-members could be assumed to be complete, in which case no "missing" end-member should be estimated (its uncertainties would anyhow be very large).

Whether the set of measured end-members is complete or not, the fitted slopes in nonzero-intercept regressions like Equations 8 and 9 depend only on correlations, and thus can be used to estimate mixing fractions in systems that have been corrupted by (constant) fractionation or measurement bias, even if unknown. By contrast, conventional mixing models typically must assume that fractionation and measurement biases are precisely known, or are zero.

5. Mixing Fractions for Systems With Memory

In many applications, tracers are sampled infrequently enough relative to the system's residence time (e.g., monthly sampling of tree xylem) that the mixture is completely replaced between samplings, consistent with the assumptions underlying both conventional end-member mixing analysis and Equations 4–7 above. In other applications, however, the sampling frequency may be high enough relative to the system's residence time (e.g., weekly sampling of soil water or groundwater) that the mixture at each time step may partly consist of carryover from the previous time step.

Conventional end-member mixing cannot account for these memory effects, but in EEMMA this is straightforward. In incomplete mixing models, where one or more sources are unmeasured, the regression model of Equation 9 can be extended by adding a term for the lagged mixture,

$$m_i = \sum_{j=1}^{k} v_{i,j}\beta_j + \beta_m m_{i-1} + \alpha + e_i, \tag{12}$$

where each mixing fraction f_j is estimated by $\beta_j/(1-\beta_m)$, the regression intercept α estimates $(1-\Sigma f)\left(\overline{\tilde{u}}\right)$, and the residual term e_i subsumes $(1-\Sigma f)\,u_i'$ plus any error in m_i (as in Section 4 above, any unmeasured sources must not be strongly correlated with any measured ones).

Alternatively, if all end-members have been sampled, a system with memory can be described by the zero-intercept regression,

$$(m_i - m_{i-1}) = \beta_1(v_{i,1} - m_{i-1}) + \cdots + \beta_k\,(v_{i,k} - m_{i-1}), \tag{13}$$

where the coefficient of the lagged mixture is estimated by $\beta_m = 1 - \sum\beta_j$ and each mixing fraction is estimated by $f_j = \beta_j/(1-\beta_m) = \beta_j/\sum\beta_j$.

The memory coefficient β_m can be considered as the "old water fraction," that is, the fraction of the mixture that is inherited from the previous time step. It can be used to infer the residence time τ of the mixture as $\tau = -\Delta t/ln(\beta_m)$, where Δt is the time interval between tracer samples. The individual β_j coefficients can be considered "new water fractions" in the sense of ensemble hydrograph separation (see Sections 4 and 5.4 of Kirchner, 2019); they quantify the fraction of the mixture originating from source j during each time step (averaged over all time steps). The mixing fractions f_j, by contrast, quantify the fraction of the mixture originating from source j at any time in the past.

As the benchmark tests in Figure S3 in Supporting Information S1 illustrate, memory effects in the mixture undermine the reliability of mixing fractions inferred from individual tracer measurements in conventional mixing models (blue dots in Figures S3b, S3e, and S3h in Supporting Information S1), because Equations 1–3 no longer hold. Nonetheless EEMMA can reliably estimate the average mixing fractions of each source, as well as the carryover fraction β_m, even if there is substantial overlap between the end-members (Figures S3c, S3f, S3i, and S4 in Supporting Information S1).

6. Accounting for Uncertainties in End-Members

The benchmark tests presented above have assumed that errors in the end-members themselves can be ignored. But if sampling and measurement errors contribute substantially to the variability of the end-members, they should be taken into account. Random errors in explanatory variables (e.g., end-members) inflate their variance,

typically creating a downward "attenuation bias" in their regression slopes (Carroll et al., 2006). Other cases are less simple. For example, in Equation 4, errors in v_1 will lead to a conventional attenuation bias in the mixing fraction f_1 but errors in v_2 will bias f_1 upward instead, because v_2 is shared between the left and right sides (Figure S5 in Supporting Information S1). The biases that arise in multiple regressions will depend on how the explanatory variables (e.g., the right-hand sides of Equations 7 and 9) are correlated among themselves, as well their correlations with the response variables on the left-hand sides, and on any correlations among their errors. (Analogous biases can also arise in conventional end-member mixing models, but have not been addressed in the literature.) Correcting for the biases in such "measurement error models" (also called "errors-in-variables problems") remains an open problem in statistics (Carroll et al., 2006). As described in Supporting Information S1, an approach first presented by Fuller (1987) can be adapted to correct these biases in the context of EEMMA, as confirmed by benchmark tests (Figures S5 and S6 in Supporting Information S1). An implementation of this approach is available as an option in the R script eemma.R (see Data Availability Statement).

7. Concluding Remarks

The approach outlined above will not work in some systems, such as those where only single measurements are available, or where measurements can't be grouped into sets or time steps. This approach will also be unnecessary whenever the restrictive assumptions of conventional mixing models are met. But it may find wide application to stable isotope tracers, which often exhibit overlapping means, rendering conventional methods infeasible. Stable isotopes are conservative to the extent that they do not undergo fractionation, but conventional mixing models require end-members with substantially different means (which thus, paradoxically, must have undergone substantial fractionation). By quantifying end-members that have overlapping means (Sections 1–3), EEMMA enables new applications for stable isotopes, which are the best available tracers for many environmental systems. EEMMA also allows a single stable isotope tracer to quantify more than two end-members (Section 2), reducing the urge to use non-conservative tracers (e.g., electrical conductivity or reactive solutes) to quantify mixing models that have many sources. And by allowing sources to be quantified even when some end-members are unknown or unmeasured, EEMMA further expands the scope of mixing models that can be solved.

EEMMA is amenable to several straightforward extensions. It could be generalized to treat multiple tracers simultaneously, thus expanding the number of end members that can be quantified from a fixed number of sampling times. Weighted regressions could be used to give greater influence to time steps when fluxes are highest (e.g., atmospheric water vapor isotopes could be weighted by humidity). If outliers are suspected, their influence could be limited using robust estimation methods such as Iteratively Reweighted Least Squares. In ill-posed systems in which end-members cannot be uniquely identified, Tikhonov regularization methods could be used to guarantee a stable approximate solution. Non-negative Least Squares could be employed to guarantee that the estimated mixing fractions are positive (at the risk of introducing an upward bias to small and uncertain values). Bayesian methods could be used to incorporate prior information (at the risk of just reinforcing one's preconceptions if the data themselves do not provide a strong constraint). Last but not least, results from EEMMA could be combined with flux information to infer how each end-member is partitioned among its different fates, thus determining where each end-member goes rather than where the mixture comes from (the end-member splitting approach; Kirchner & Allen, 2020). But even without such enhancements, by exploiting fluctuation information contained in tracer time series, EEMMA substantially expands the scope of feasible mixing models.

Data Availability Statement

An R script that performs EEMMA (eemma.R), along with an example data set and a brief user's guide, are available at https://www.doi.org/10.16904/envidat.410 (Kirchner, 2023). Version 1.0, build 2023.03.27, was used in this paper.

Acknowledgments
This analysis was motivated by collaborative research with Qin Liu of Tianjin University concerning isotope tracers of plant water uptake. I also thank Scott Jasechko, Thom Bogaard, Paolo Benettin, Marius Floriancic, Fabian Bernhard, Scott Allen, and Harsh Beria for helpful discussions.

References

Barthold, F. K., Tyralla, C., Schneider, K., Vache, K. B., Frede, H. G., & Breuer, L. (2011). How many tracers do we need for end member mixing analysis (EMMA)? A sensitivity analysis. *Water Resources Research*, *47*(8), W08519. https://doi.org/10.1029/2011WR010604

Beria, H., Larsen, J. R., Michelon, A., Ceperley, N. C., & Schaefli, B. (2020). HydroMix v1.0: A new Bayesian mixing framework for attributing uncertain hydrological sources. *Geoscientific Model Development*, *13*(5), 2433–2450. https://doi.org/10.5194/gmd-13-2433-2020

Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: A modern perspective*. Chapman & Hall/CRC.

Christophersen, N., Neal, C., Hooper, R. P., Vogt, R. D., & Andersen, S. (1990). Modelling streamwater chemistry as a mixture of soilwater end-members—A step towards second-generation acidification models. *Journal of Hydrology*, *116*(1–4), 307–320. https://doi.org/10.1016/0022-1694(90)90130-P

Fuller, W. A. (1987). *Measurement error models*. Wiley.

Gessler, A., Ferrio, J. P., Hommel, R., Treydte, K., Werner, R. A., & Monson, R. K. (2014). Stable isotopes in tree rings: Towards a mechanistic understanding of isotope fractionation and mixing processes from the leaves to the wood. *Tree Physiology*, *34*(8), 796–818. https://doi.org/10.1093/treephys/tpu040

Jiskra, M., Heimbürger-Boavida, L.-E., Desgranges, M.-M., Petrova, M. V., Dufour, A., Ferreira-Araujo, B., et al. (2021). Mercury stable isotopes constrain atmospheric sources to the ocean. *Nature*, *597*(7878), 678–682. https://doi.org/10.1038/s41586-021-03859-8

Kirchner, J. W. (2019). Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: Theory and benchmark tests. *Hydrology and Earth System Sciences*, *23*(1), 303–349. https://doi.org/10.5194/hess-23-303-2019

Kirchner, J. W. (2023). eemma.R, an R script for ensemble end-member mixing analysis [Software]. EnviDat. https://www.doi.org/10.16904/envidat.410

Kirchner, J. W., & Allen, S. T. (2020). Seasonal partitioning of precipitation between streamflow and evapotranspiration, inferred from end-member splitting analysis. *Hydrology and Earth System Sciences*, *24*(1), 17–34. https://doi.org/10.5194/hess-24-17-2020

Klaus, J., & McDonnell, J. J. (2013). Hydrograph separation using stable isotopes: Review and evaluation. *Journal of Hydrology*, *505*, 47–64. https://doi.org/10.1016/j.jhydrol.2013.09.006

Layman, C. A., Araujo, M. S., Boucek, R., Hammerschlag-Peyer, C. M., Harrison, E., Jud, Z. R., et al. (2012). Applying stable isotopes to examine food-web structure: An overview of analytical tools. *Biological Reviews*, *87*(3), 545–562. https://doi.org/10.1111/j.1469-185X.2011.00208.x

Stock, B. C., Jackson, A. L., Ward, E. J., Parnell, A. C., Phillips, D. L., & Semmens, B. X. (2018). Analyzing mixing systems using a new generation of Bayesian tracer mixing models. *PeerJ*, *6*, e5096. https://doi.org/10.7717/peerj.5096

Tetzlaff, D., Buttle, J., Carey, S. K., McGuire, K., Laudon, H., & Soulsby, C. (2014). Tracer-based assessment of flow paths, storage and runoff generation in northern catchments: A review. *Hydrological Processes*, *29*(16), 3475–3490. https://doi.org/10.1002/hyp.10412

References From the Supporting Information

Genereux, D. (1998). Quantifying uncertainty in tracer-based hydrograph separations. *Water Resources Research*, *34*(4), 915–919. https://doi.org/10.1029/98WR00010

Phillips, D. L., & Gregg, J. W. (2001). Uncertainty in source partitioning using stable isotopes. *Oecologia*, *127*(2), 171–179. https://doi.org/10.1007/s004420000578