https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

Improved endmember mixing analysis (EMMA): Application to a snowmelt-dominated stream in northern Utah

Alyssa N. Thompson[1], Barry R. Bickmore[1], Gregory T. Carling[1], Emily J. Evans[2], Stephen T. Nelson[1], Joshua J. LeMonte[1], Kevin A. Rey[1], Diego P. Fernandez[3], Kendra L. Caskey[1]

[1]Department of Geological Sciences, Brigham Young University, Provo, 84602, USA
[2]Department of Mathematics, Brigham Young University, Provo, 84602, USA
[3]Department of Geology & Geophysics, University of Utah, Salt Lake City, 84112, USA

Correspondence to: Greg Carling (greg.carling@byu.edu)

Abstract. An endmember mixing analysis (EMMA) is a sophisticated hydrograph separation technique used to determine the primary water sources in a watershed and estimate their respective input over time. In a traditional EMMA approach, a principal component analysis (PCA) is used to identify endmember composition, and the retained principal component (PC) scores are used to calculate the fractional contributions of each endmember. This approach is based on the idea that the reduced dimensionality of the endmember data in just a handful of PCs contains the most useful information. While this approach does simplify the mixing calculation, it limits potential model complexity. We show that calculating endmember contributions using the original water chemistry data (tracer space) results in a more simplified and uniform approach than performing the calculation in PC-defined subspace. Additionally, we demonstrate an iterative approach to selecting the tracers and endmembers to create a more complex (and more representative) model. We applied EMMA to the upper Provo River watershed (262 km^2), a snowmelt-dominated catchment in northern Utah, to test some potential improvements in the method. Five endmembers (quartzite groundwater, carbonate groundwater, mineral soil water, organic soil water, and snow) were identified for the watershed and differentiated using seven tracers ($\delta^{18}O$, δ^2H, HCO_3^-, Si, Mg^{2+}, K^+, and Ca^{2+}). We applied this approach in a well-defined workflow implemented in EMMALAB, a software application designed to perform EMMA on one or more stream locations in a catchment. The analysis showed that snow was the dominant endmember during spring runoff, contributing 38% of flow on average, while quartzite groundwater contributed 60% during baseflow. The iterative analysis for selecting endmembers and tracers is easily implemented through EMMALAB, allowing for a uniform and simplified approach to apply the complex mathematics behind EMMA for more accurate hydrograph separation calculations.

1 Introduction

Hydrograph separation is a critical tool for quantifying the contributions of surface and subsurface water sources to streams, which is especially important for forecasting the evolution of water resources and water quality due to climate change and

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

human activities. As we attempt to characterize larger, more complex watersheds, it will also be necessary to apply hydrographic separation methods that are capable of handling added complexity.

Endmember mixing analysis (EMMA) is a relatively sophisticated hydrograph separation method (Christophersen and Hooper, 1992) in which endmember water sources (e.g., groundwater, precipitation, snow, soil water) are defined using conservative chemical or isotopic tracers (e.g., Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Si, Cl^{-}, HCO_3^{-}, NO_3^{-}, $\delta^{18}O$, $\delta^{2}H$, and electrical conductivity) (Christophersen et al., 1990a; Hooper, 2003; Barthold et al., 2011). EMMA traditionally relies on a principal component analysis (PCA) to reduce the dimensionality of the data, making it possible to visualize and identify clusters of data, and to use a greater variety of data features in the analysis than was previously the case with isotopic hydrograph separation methods (Christophersen and Hooper, 1992). Visualizing the data in a dimensionally reduced space allows for identifying potential endmembers that circumscribe the stream chemistry data. The extent of mixing between the various endmembers is calculated by solving a system of linear equations, typically involving the dimensionally reduced data, to obtain the fractional contributions of each endmember (Birkel et al., 2021; Jacobs et al., 2018; Lukens et al., 2022; Foks et al., 2018; Bugaets et al., 2023; Christophersen and Hooper, 1992).

Given that EMMA uses more types of data than most hydrograph separation methods, it is a good starting point for hydrological analysis of relatively complex watersheds. Our impression, however, is that there are two barriers to maximizing its usefulness: 1) there has been no standard software package to streamline the process, and 2) the mathematical approach typically used in the mixing calculation is not optimal for expanding model complexity.

EMMA model calculations have to this point usually been programmed by individual researchers and so may have been underutilized by researchers who lack the mathematical or computational background to set up the calculations correctly (Chaves et al., 2008; He et al., 2020; Petermann et al., 2018). EMMA also requires several subjective choices throughout the process that can be difficult and time-consuming to iterate upon without a standardized approach. These choices primarily consist of tracer selection, endmember selection, and the number of principal components (PCs) to retain. As a result of this difficulty, some studies offer flawed justification for tracer selection (Chisola et al., 2022; Guinn et al., 2010; Liu et al., 2004; Lund et al., 2023) and endmember selection (Montagud et al., 2021).

The standard method of performing the mixing calculations in dimensionally reduced space is problematic for two reasons. First, the number of dimensions used in the calculation places a limit on the number of endmembers that can be included, in turn limiting the hydrological complexity that can be modelled. Second, while dimensional reduction via PCA is indispensable for visualizing the relationship between potential endmembers and stream chemistry data, there is no compelling reason to perform the mixing calculation in dimensionally reduced space even when only a few endmembers are needed. Whereas some assume that the reduced dimensionality removes unimportant "noise" from the data, there is no way to know in advance whether any of the discarded information will turn out to be important (Lichtert and Verbeeck, 2013).

In this contribution, we critically evaluate previous EMMA approaches and introduce a new software package (EMMALAB) to facilitate a rapid, iterative modelling process that yields more reliable results and can accommodate maximum complexity. To illustrate, we apply the method to the upper Provo River, a snowmelt-dominated perennial stream located in the Uinta

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

Mountains of northern Utah, USA. The Provo River contributes to the drinking water supply for over half of Utah's population. The upper Provo River watershed covers a significantly larger area (262 km^2) than is typically analysed (1 – 100 km^2) via EMMA (Durand and Torres, 1996; Cuoco et al., 2021; Guinn et al., 2010; Barthold et al., 2010; Lukens et al., 2022; Christophersen et al., 1990b; Christophersen and Hooper, 1992), but the lithology is relatively simple, making it an ideal study site for our purposes.

2 Materials and Methods

An endmember mixing analysis (EMMA) is a tracer-based hydrograph separation technique carried out via the following steps proposed by Christophersen et al. (1990a) and (Christophersen and Hooper, 1992): (1) measure chemistry of stream water and water sources (e.g., groundwater, precipitation, soil water, etc.) that could be used as potential endmembers for hydrograph separation; (2) determine which solutes or isotopes behave conservatively for use as tracers; (3) identify which potential endmembers can be plausibly applied to the model using a principal components analysis; (4) optimize the mixing model to calculate the fractional endmember contributions to the stream; and (5) perform an error analysis of the endmember datasets. In the following subsections, we describe the Provo River watershed, where we applied EMMA, and provide further details on how we have modified established EMMA methods to simplify and improve mixing estimates.

2.1 Upper Provo River watershed site description

The upper Provo River watershed is located in the southwestern Uinta Mountains of northern Utah (Figure 1). The watershed is dominantly fed by high elevation snowmelt and receives diverted water from the Duchesne River as part of the Provo River Project (U.S. Bureau of Reclamation, 1958).The watershed above Soapstone, including the upper Duchesne River watershed above the diversion, covers 262 km^2.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

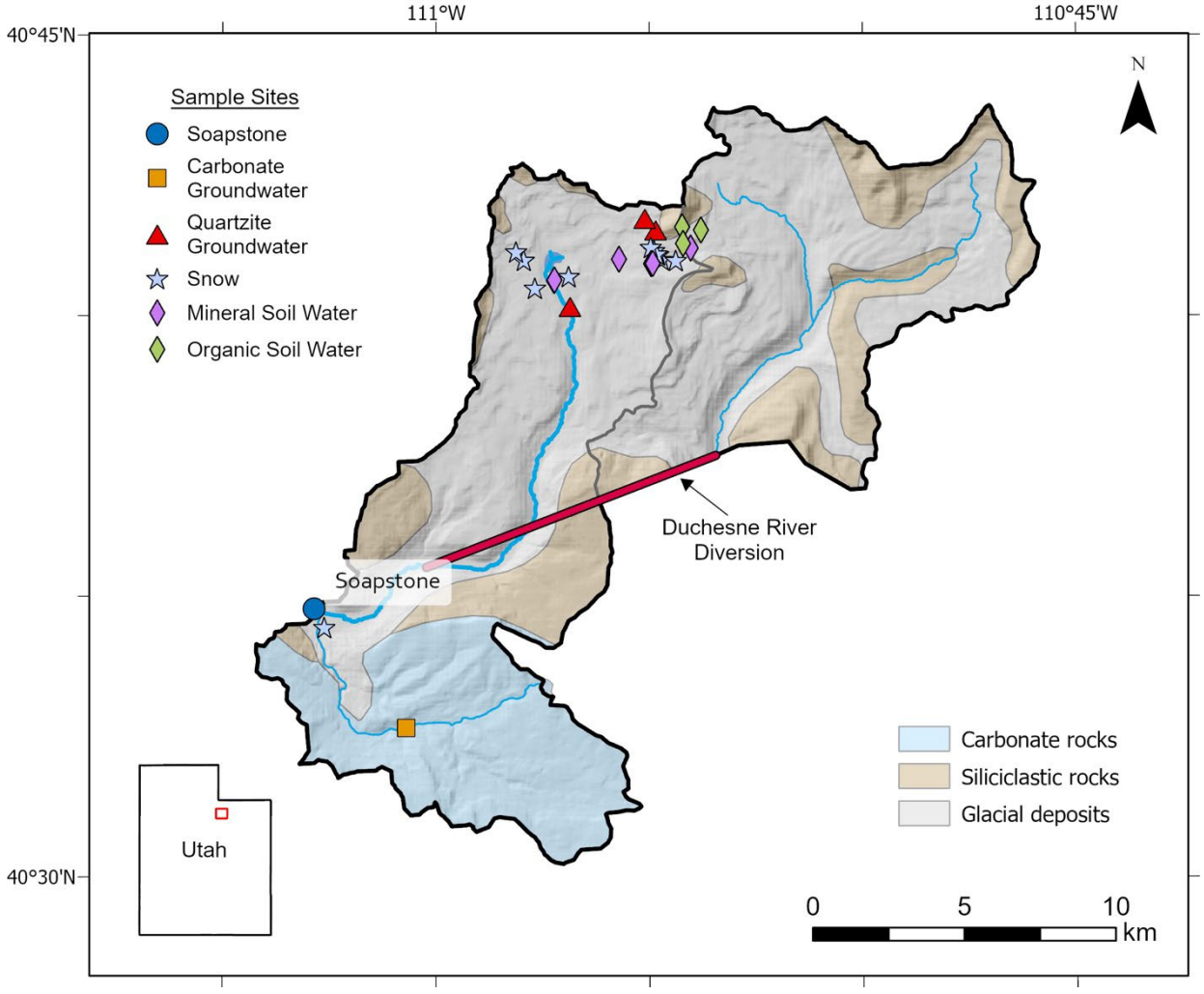


Figure 1: Delineated watershed boundaries for the upper Provo and Duchesne rivers in northern Utah, USA. The river sampling location at Soapstone was selected at the transition between underlying siliciclastic and carbonate rocks. Endmember sampling sites were exclusively collected within the upper Provo River watershed. Geologic map modified after Bryant (2003).

The upper Provo River extends ~19 km with a vertical relief of 540 m, from 2900 m asl at the headwaters at Trial Lake to 2360 m asl at Soapstone. The discharge for Soapstone was estimated using available datasets. Specifically, the measured Duchesne diversion discharge was subtracted from the US Geological Survey-measured discharge at Woodland (located ~15 km downstream of Soapstone on the Provo River) then multiplied by the ratio of the area of Soapstone and Woodland watersheds. The Duchesne diversion discharge was then added back to this value to create the final estimated discharge values at Soapstone from 2016 to 2022 (Figure 2).

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

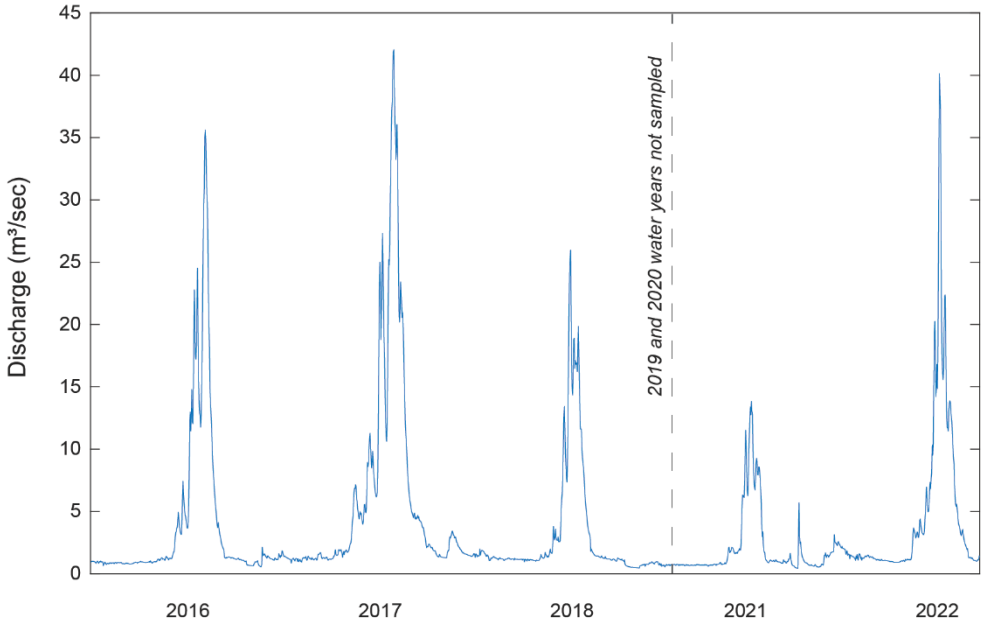


Figure 2: Time-series daily discharge rates for Soapstone from 2016 – 2018 and 2021- 2022.

The geology of the watershed primarily consists of weakly metamorphosed Neoproterozoic siliciclastic rocks with minor amounts of shale, overlain by glacial and other Quaternary deposits (Figure 1). The soil depth is typically ~1 m (Munroe et al., 2015), enriched by dust in the A soil horizons with a different composition than the underlying lithology (Munroe et al., 2020). The soil types of the area are mostly dominated by inceptisols, with some alfisols, mollisols, and histosols. Vegetation primarily consists of Lodgepole pines and Engleman spruces with an understory of huckleberry, sedges, and moss. The study area is mostly undeveloped except for the Mirror Lake Highway and several campgrounds.

2.2 Stream and endmember water sample collection

Water samples were collected from the Provo River at Soapstone during water years 2016 through 2018 and 2021 through 2022. Samples were collected once or twice per month, with more intensive weekly or twice-weekly sampling during the snowmelt runoff period (April – June) each year. A total of 106 stream samples were collected over the five-year period. Potential endmember samples were collected between 2012 and 2022. Snowpack endmember samples (n = 22) were collected prior to spring snowmelt at maximum accumulation. Samples were collected by digging snow pits and taking a full-depth snow core using an acid-washed plastic tube. Samples from the plastic tube were deposited into clean FLPE bottles. The bottles were melted in the lab, then sub-sampled for chemical analysis.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

Soil water endmember samples for two different soil water endmembers, mineral soil water and organic soil water, (n = 6 and n = 3) were collected during spring snowmelt to allow for maximum saturation of the soils. The mineral soil water samples were collected from an inceptisol, and the organic soil water samples were collected from an alfisol containing a litter layer. Both soil types are a similar geomorphology, consisting of deep, well-drained permeable soils on mountain moraines with a relatively thin soil layer. Samples were collected by using a small shovel to collect saturated soil and placed into a plastic bag. Soils were transferred into 50 mL centrifuged tubes and centrifuged at 6000 rpm for 20 minutes. The supernatant was decanted and sub-sampled for various analyses.

Quartzite (n = 9) and carbonate (n = 5) groundwater endmember samples were collected from springs throughout the watershed during summer and fall (between July and October) to collect representative chemistry with limited mixing with snowmelt. Springs discharging from siliciclastic and carbonate rocks were classified as quartzite and carbonate groundwater, respectively. Although no quartzite or carbonate groundwater springs were sampled in the Duchesne watershed, the chemistry of these endmembers was assumed to be uniform across the two watersheds.

Stream and endmember samples were analysed for major cations and trace elements, major anions, bicarbonate, and $\delta^{18}O$ and $\delta^{2}H$ isotopes. Samples for major cations and trace elements were filtered with a 0.45 µm PES syringe filter into an acid-washed 60 mL LDPE bottle and acidified to 2.4% v/v TMG nitric acid. Samples for major anions and bicarbonate were collected in a 1 L plastic bottle and further processed in the lab. Water samples for $\delta^{18}O$ and $\delta^{2}H$ analysis were collected in amber glass bottles with Polyseal caps to prevent evaporation. All samples were refrigerated until analysis.

2.3 Laboratory analysis

Water samples were collected for analysis of major cations and trace elements using an Agilent 7500ce quadrupole inductively coupled plasma mass spectrometer (ICP-MS). The following elements were measured for all stream and endmember samples: Al, As, B, Ba, Be, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Gd, Ho, K, La, Li, Lu, Mg, Mn, Mo, Na, Nd, Ni, Pb, Pr, Rb, Sb, Sc, Se, Sm, Sr, Tb, Th, Tl, U, V, Y, Yb, and Zn. Further descriptions of the methods are given in Checketts et al. (2020). Inductively coupled plasma optical emissions spectroscopy (ICP-OES) was additionally used to analyse for Si as the detection limits through ICP-MS were insufficient to accurately determine Si concentrations. Major anion (F^-, Cl^-, NO_3-, and SO_4^{2-}) concentrations were analysed using a Dionex ICS-90 ion chromatograph (IC). HCO_3^- concentrations were analysed on unfiltered samples by acid titration using a Mettler Toledo DL50 titrator on the samples collected through 2018 and titrated by hand using a Hach alkalinity test kit on the samples collected after 2020. Alkalinity values were assumed to be equal to bicarbonate concentrations. Water stable isotope ($\delta^{18}O$ and $\delta^{2}H$) ratios in water were analysed using a Los Gatos Research Liquid Water Isotope Analyzer with results reported relative to Vienna Standard Mean Ocean Water (VSMOW).

To evaluate analytical errors in the major anion and major cations datasets, we calculated the charge balance error for each sample. Most samples had acceptable charge balance error using the raw data. However, any sample that was outside of +/- 5% charge balance error was adjusted by varying the bicarbonate concentration. Bicarbonate values were determined to be the

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

solute concentrations with the most uncertainty and were adjusted in 50 out of 106 Soapstone samples, with an average adjustment of 9.5 mg/L. Of the total adjustments made, 88% of the samples were adjusted from samples collected during 2020 and later. This is likely due to the change in methods in measuring alkalinity beginning in 2020. The adjustments to the 2020-2022 samples resulted in values that were consistent with concentrations measured in the earlier samples.

2.4 Tracer selection

To evaluate which tracers could be used in the mixing analysis, we investigated the potential conservative behaviour of isotopes and solutes in the full dataset. Isotopes and solute concentrations are considered conservative if they do not undergo isotopic fractionation or participate in chemical reactions, respectively, which can sometimes be observed by linear relationships in bivariate plots (Hooper, 2003). Although these linear relationships may not always suggest conservative behaviour or identify all conservative tracers within the dataset, it is a useful tool to create an initial set of tracers from which to test the model. $\delta^{18}O$ and $\delta^{2}H$ values can be assumed to be conservative, even if they do not exhibit linear relationships with respect to the solute tracers, because $\delta^{18}O$ and $\delta^{2}H$ track elevation and meteoric precipitation rather than water-rock interactions (Clark and Fritz, 1997).

We used a stepwise approach to determine the most reasonable conservative tracers to include in our analysis. The conservative behaviour of potential tracers was determined from an array of 49 elements and two stable isotopes, $\delta^{18}O$ and $\delta^{2}H$. As many conservative elements as possible were retained to maximally constrain the mixing models. In our case, bivariate plots of solutes and isotopes with strong linear relationships ($R > 0.80$) were assumed to behave conservatively. As a further step, common conservative tracers that did not plot linearly on the bivariate plots were iteratively included in the model in a stepwise approach and then their predicted values were assessed for accuracy. If the tracer predicted values had a low RMSE on the predicted vs. observed plots, then it was assumed that the tracer was conservative and was thus retained in the final model. In EMMALAB, this is streamlined by allowing the user to iteratively include or exclude any tracer into the model and rerun the mixing calculation, allowing for quick and thorough analysis of the potential tracers for the model. In this way, the user is free to assess any tracer they may want to include in the model and increase the chance of establishing a useful tracer set. A key principle of an endmember mixing analysis is for the endmembers to have distinct chemical or isotopic composition (Cuoco et al., 2021; James and Roulet, 2006). By including all conservative solutes or isotopes in the analysis, tracers most distinctly representative of each endmember are more likely to be recognized.

2.5 Data formatting and standardization

In preparation for applying EMMA, both the stream and endmember data must be properly organized and standardized. Here we describe the mathematical details.

Let **R** be an $m \times n$ matrix where m is the number of stream water samples and n is the number of tracers. We will refer to each column of **R** as vector $\vec{r}_j$, which is the collection of tracer j values in the m stream water samples. The individual elements of

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

matrix **R**, r_{ij}, are the concentrations (or isotopic ratios) of tracer *j* in sample *i*. Prior to the PCA, concentrations in the matrix **R** were standardized using *z*-scoring, where the mean of each column of the solute data is subtracted from the values in the column, and then divided by the standard deviation of that column. This allows each column to have a mean of zero and a standard deviation of 1, ensuring that the subsequent analysis is not dominated by variables with large absolute values (Miesch, 1980; Christophersen and Hooper, 1992). The resulting matrix, $\mathbf{R}_z$, is then composed of column vectors $\vec{r}_{z_{ij}}$, which are the *z*-scored values of tracer *j* in sample *i*.

$$\mathbf{R} = [\vec{r}_1 \quad \dots \quad \vec{r}_j \quad \dots \quad \vec{r}_n] = \begin{bmatrix} r_{11} & \cdots & r_{1j} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{i1} & \cdots & r_{ij} & \cdots & r_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mj} & \cdots & r_{mn} \end{bmatrix} \tag{1}$$

$$r_{z_{ij}} = \frac{r_{ij} - \mathrm{m}_{\vec{r}_j}}{\mathrm{s}_{\vec{r}_j}} \tag{2}$$

Prior to performing EMMA, each dataset of individual endmembers should be averaged together to create a single sample that represents the single endmember for that area. Each of these averaged endmembers should then be arranged in a *d*×*n* matrix (**M**), where *d* is the number of averaged potential endmembers and *n* is the number of tracers. Prior to calculating the principal component (PC) scores for the endmembers, concentrations are first standardized using the mean and standard deviation of the stream sample tracers. The new matrix of standardized endmember values is $\mathbf{M}_z$.

$$\mathbf{M} = [\vec{m}_1 \quad \dots \quad \vec{m}_j \quad \dots \quad \vec{m}_n] = \begin{bmatrix} m_{11} & \cdots & m_{1j} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{i1} & \cdots & m_{ij} & \cdots & m_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{d1} & \cdots & m_{dj} & \cdots & m_{dn} \end{bmatrix} \tag{3}$$

$$m_{z_{ij}} = \frac{m_{ij} - \mathrm{m}_{\vec{r}_j}}{\mathrm{s}_{\vec{r}_j}} \tag{4}$$

In EMMALAB, this process is automated when the user inputs their data through the given example spreadsheet. The chemical isotopic data is taken from the input spreadsheet and the necessary calculations to set up the mixing model are performed automatically. Details on the specific coding behind these processes are detailed in DataEMMA.m.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

2.6 Endmember selection

A principal component analysis (PCA) was performed on the standardized stream data $\mathbf{R}_z$ to determine which potential endmembers to use in EMMA. A PCA is a dimensional reduction technique used to reduce the number of variables to a more manageable level. In a PCA, a new set of n variables (principal components or PCs), composed of linear combinations of the *n* original variables, is created by rotating, stretching and shrinking the original axes to maximize the amount of data variance explained by the fewest number of PCs. The PCs are then ranked in descending order of how much data variance they describe, so that the first principal component (PC_1) describes the most variance and PC_n describes the least. The PCs contain the same information as the original data, but if some of the later PCs are discarded, less information is lost than if the same number of the original variables were discarded.

In our case, the PCA was done using the MATLAB® *pca* function, which accepts $\mathbf{R}_z$ as input and produces several outputs, including the following. First, it produces an *n*×*n* coefficient matrix, $\mathbf{C_R}$, composed of column vectors representing the coefficients of linear equations that transform the original data coordinates into principal component space (Table 1).

Table 1: Values from the coefficient matrix for the river data. Three PCs were used to visually observe the river data and select endmembers, with particular weight being on HCO_3^-, $Mg2^+$ and Ca^{2+} for PC1, $\delta^{18}O$ and δ^2H for PC2, and Si and K^+ for PC3.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
$\delta^{18}O$	-0.12	**0.66**	0.23	0.07	-0.69	-0.12	0.10
δ^2H	-0.12	**0.68**	0.12	-0.02	0.72	-0.03	0.02
HCO_3^-	**0.54**	0.16	-0.23	-0.10	-0.03	-0.50	-0.61
Si	0.11	-0.15	**0.70**	-0.68	0.02	-0.07	-0.04
Mg^{2+}	**0.55**	0.17	-0.10	0.08	-0.05	0.79	-0.18
K^+	0.24	-0.18	**0.58**	0.70	0.11	-0.25	0.04
Ca^{2+}	**0.56**	0.07	-0.19	-0.14	0.03	--0.19	0.77

A matrix of principal component scores ($\mathbf{R}_{PC}$), contains the original data values transformed in terms of the new principal component space, and is produced by multiplying the matrix $\mathbf{R}_z$ by $\mathbf{C_R}$.

$$\mathbf{R}_{PC} = \mathbf{R}_z * \mathbf{C_R} \quad (5)$$

Next, the function produces an *n*×1 vector ($\vec{e}_{\mathbf{R_{PC}}}$) of eigenvalues of the covariance matrix of $\mathbf{R}_z$, where each value represents the amount of variance explained by the corresponding principal component. An eigenvalue of one (for *z*-scored data) means that the corresponding principal component explains the same amount of variance as any one of the original variables, whereas any value higher than one means that it explains more variance than the original variables (Davis, 2002). The principal

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

components are arranged in descending order of the amount of variance they explain, so in order to visually represent the data, the first few principal components are retained and the rest are discarded to reduce the dimensionality of the data while retaining as much of the total variance as possible.

There are multiple ways to determine the number of principal components to retain for the analysis. Most commonly, an arbitrary threshold is set for the percent of variance that should be explained by the retained PCs, typically varying between 70 – 90% (Montagud et al., 2021; Jolliffe and Cadima, 2016). Similarly, the rule of one suggests retaining all PCs with eigenvalues equal to or greater than 1 (Joreskog et al., 1976; Barthold et al., 2010; James and Roulet, 2006; Hooper, 2003), meaning that they explain at least as much variance as the original variables. Another method includes creating a scree plot, which is especially useful if the user intends for the model to explain a certain fraction of the total variance in their analysis. A scree plot often shows a steep decline in the magnitude of eigenvalues as the number of PCs increases, with the plot eventually flattening out beyond a certain point. This point is known as the "elbow" of the plot, and its location indicates the number of PCs to retain in the analysis (Cattell, 1966). All these methods for determining the number of PCs to retain are included in EMMALAB, including a display of the variance explained by each PC, eigenvalues and a scree plot.

Once a subset of the PCs is retained for further use, the space defined by these PCs is called U-space (Christophersen and Hooper, 1992) and the data is projected on U-space for visual examination to aid in endmember selection. The minimum number of endmembers to include in the analysis is the number of retained PCs plus one (Christophersen and Hooper, 1992). In terms of PC space (U-space), it takes a minimum of U+1 points to describe a convex hull in U-dimensional space, where U is the number of retained PCs. That is, it takes at least three points to circumscribe a cloud of data in two-dimensional space, and four points to circumscribe a cloud of data in three-dimensional space. If more than three PCs are retained, selecting the appropriate number of endmembers becomes more challenging as only three-dimensions can be visually represented at once (Davis, 2002). Because U-space must be defined by the stream data and not the endmember data, PC scores for the stream endmembers are calculated using the coefficient matrix $\mathbf{C_R}$. The PC scores for the endmembers are calculated from the standardized endmember values $\mathbf{M}_z$ and multiplied by the coefficient matrix $\mathbf{C_R}$.

$$\mathbf{M}_{PC} = \mathbf{M}_z * \mathbf{C_R} \tag{6}$$

The selected PC scores for the stream water and endmembers samples are projected into U-space to determine which set of endmembers adequately circumscribe the data. The selected endmembers are then treated as the primary water sources in the watershed. From this point forward, $\mathbf{M}_z$ is constructed from the standardized mean tracer values of only the selected endmembers. EMMALAB automatically transforms the endmember data in terms of the PC stream data and the code can be found on the class definition, DataEMMA.m. The primary purpose of the PCA is to select endmembers for the analysis, and we propose that the mixing calculation should be performed in the original tracer space rather than in U-space, as described below.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

2.7 Mixing calculation

The primary purpose of the PCA in EMMA is to select plausible endmembers for the mixing analysis, but initial studies establishing EMMA also performed the mixing calculations in U-space (defined by the subset of retained PCs) (Christophersen and Hooper, 1992). We argue that there is no compelling reason to adopt this strategy, and that performing the mixing calculations in standardized tracer space is preferable.

Mixing models are most commonly used to determine the fractional input from multiple sources (endmembers) to the main area of interest, like a stream or sediment sample. A general way of expressing this mixing is in the form of a summation (Albarède, 2009), where C_0^j is the concentration of tracer j in the stream sample, f_w is the fractional contribution from endmember w, C_w^j is the concentration of tracer j in endmember w, and d is the number of endmembers. This solution adheres to the principle of conservation of mass by assuming that the concentrations of all defined inputs or endmembers into the system will fully explain the concentrations in the mixture. This introduces the following set of constraints:

$$\sum_{w=1}^{d} f_w = 1 \tag{7}$$

$$C_0^j = \sum_{w} f_w C_w^j \tag{8}$$

The summations may then be solved for the f_w values as a system of linear equations. This can be done via direct calculation (e.g., Gaussian elimination, etc.) if the system is critically determined (i.e., the number of variables equals the number of equations), but this does not allow one to constrain the f_w values to be non-negative, which can lead to non-physical results (Ali et al., 2010; Liu et al., 2004). Therefore, it is preferable to solve the system via optimization, which can deal with both critically determined and overdetermined (the number of variables is less than the number of equations) systems and can include various types of constraints. This defines the system of equations as a constrained, least-squares linear estimation problem to solve for the fractional contributions of each source, as described in Christophersen et al. (1990b), where n is the number of tracers j in the dataset. In any case, where $n > (d - 1)$, the result is an overdetermined system of equations.

$$\begin{aligned} 1 &= f_1 + f_2 + \dots f_d \\ C_0^j &= f_1 C_1^j + f_2 C_2^j + \cdots f_d C_d^j \\ C_0^n &= f_1 C_1^n + f_2 C_2^n + \cdots f_d C_d^n \end{aligned} \tag{9}$$

Rather than using the raw or standardized tracer concentrations in the system of linear equations, (Christophersen and Hooper, 1992) recommended using the composite concentrations in the retained PC scores of the stream and endmember samples as

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

the input. (Each PC depends on all the original variables to different degrees, which is why we refer to them as composite variables.) Typically, the user determines the number of PCs to retain for the mixing analysis (typically 2 or 3) based on the amount of variance explained by these PCs, and discards the remaining data, usually assumed to be noise. However, there is not a uniform approach to determining the number of PCs to retain, and although guidelines exist for the selection process, their application is quite subjective. Additionally, although a PC may explain only a small percentage of the variance, it may contain data useful for the mixing analysis instead of just 'noise' (Jolliffe and Cadima, 2016; Wold et al., 1987). To simplify the overall process and avoid these potential errors, the full set of standardized data should be used. This method of using the z-scored tracer concentrations to calculate endmember fractional contributions has been used in previous studies, but primarily in scenarios where the endmember compositions are unknown (Carrera et al., 2004; Liu et al., 2020; Valder et al., 2012). These studies also do not compare the mixing results when performing the calculation in U-space versus tracer space. To address this in our study, fractional contributions were calculated both using standardized stream water data and the PC scores for the stream data and compared using the optimization described in Christophersen et al. (1990a).

We set up the optimization as shown in Equations 10 – 11. The objective function minimized in our analysis is the sum of squared error, where the inputs include an n×d matrix of standardized endmember values ($\mathbf{M}_z$) and an n×1 column vector of standardized stream water concentrations from each row in matrix $\mathbf{R}_z$. The reason for using standardized values rather than tracer values is to avoid bias in the optimization because endmembers with smaller tracer concentrations would likely exert minimal influence on the optimization, negating the advantage of including as many tracers as possible. The constraints on the optimization include that the sum of the fractional contributions should be equal to 1, and the value of each contribution is constrained to $0 \leq f_w \leq 1$. This calculation is then used for each stream sample to determine the fractional contribution of each endmember for a given sampling day.

$$E_j = \left(\mathbf{M_Z} * \vec{f}\right) - \vec{\mathbf{r}}_{\mathbf{z_j}} \tag{10}$$

$$\underset{x\in[0,1]}{\arg\min}\left(\sum_{j=1}^{n} {E_j}^2\right), \qquad \text{subject to: } \sum_{w=1}^{d} f_w = 1 \tag{11}$$

In our case, the fractional contribution values in $\vec{f}$ were optimized for each stream sample using the MATLAB function fmincon. This function minimizes an objective function to determine the best values for the adjustable parameters, given a set of initial guesses and various types of constraints. Initial guesses for each endmember were set to $1/d$, where d is the number of endmembers used in the analysis. The output from fmincon is a d×1 column vector ($\vec{f}$) of estimated fractional contributions for a single stream sample, where the elements (f_w) are the individual endmember contributions in the order they were arranged in $\mathbf{M}_z$.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

$$\vec{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_d \end{pmatrix} \quad (12)$$

Although it is recommended to use the standardized tracer data in the mixing analysis, EMMALAB also includes the option to calculate the fractional contributions using either 2 PC's or 3 PC's. A tab is included within the app to compare the difference in calculated endmember fractional contributions in a line graph. The app also represents the fractional contributions calculated for each sampling day plotted in a stacked bar graph to observe how these contributions change over time. A weighted discharge plot can also be created outside of the app by exporting the calculated fractional contributions and multiplying these values by the rate of discharge measured at the time of sample collection to observe seasonal fluctuations of the relative contributions from each endmember.

2.8 Residual analysis

To check the accuracy of the model and the endmembers selected, the predicted concentrations based on the estimated fractional contributions can be compared to the actual stream chemistry for each sampling day (Foks et al., 2018). To convert the estimated fractional contributions into stream concentrations, the vector $\vec{f}$ is multiplied by the matrix of standardized endmembers, resulting in a $m \times n$ matrix of predicted standardized stream values, $\mathbf{R}_z^*$. Each standardized value is then de-standardized using the standard deviation and mean calculated for each tracer in the original matrix $\mathbf{R}$.

$$\mathbf{R}_z^* = \mathbf{M}_z * \vec{f} \quad (13)$$

$$\boldsymbol{r}_z^* = (r_{z_{ij}}^* * \mathrm{s}_{\mathbf{r}_j}) + \mu_{\mathbf{r}_j} \quad (14)$$

EMMALAB performs these calculations automatically in the script titled AnalyzeEMMA.m. Within the app, the predicted concentrations are plotted against the observed concentrations and assessed for linearity and adherence to the 1:1 line. Tracers that are under- or overpredicted by the model indicate the possible need for an additional endmember or suggest a physical or chemical process that makes them non-conservative.

2.9 Endmember error analysis

Methods for determining error used in previous studies include gaussian error propagation, bootstrapping, or Monte Carlo simulations (Vonk et al., 2010; Genereux, 1998; Bazemore et al., 1994; Xu Fei and Harman, 2022; Christophersen et al., 1990a). In our case, an error analysis was performed on the endmembers using jackknifed mean values of the endmember tracer compositions to measure error due to the variations in the endmember samples. This method was used on the original endmember datasets before they were averaged. The method of jackknifing the mean includes generating a distribution of

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

mean values by systematically excluding a single observation and averaging the remaining observations (Trauth, 2021). As mentioned above, the endmembers used in the model are each composed of a group of samples that are averaged together to represent the endmember. If each group of samples for an individual endmember has m samples, then a new representative averaged endmember is created by removing one sample from the dataset and averaging the remaining m-1 samples. This is repeated m times, creating a new matrix the same dimensions as the initial matrix of endmember samples, but now containing a variety of averaged endmember values for the individual endmember. This is repeated for each endmember used in the analysis, and the averaged groups calculated for each endmember are arranged in every possible combination and used in the optimization for every sample collected. If a dataset has four endmembers and 10 samples collected for each endmember, there would be 10,000 iterations of different averaged endmember combinations used to calculate the fractional contribution of each endmember for each sample. The resulting mean and standard deviation for the percent contribution of all iterations for each sampling day are then used to quantify the extent of error due to endmember sampling. Within EMMALAB, the MATLAB function *jackknife* is used to calculate the jackknifed mean values for each endmember. These values can then be plotted as error bars on a line graph around the endmember fractional contribution values over time.

3 Results and Discussion

Due to the number of subjective choices that must be made in a typical EMMA application, the ability to iterate over several model permutations is critical. However, this can be very time-consuming if every change must be individually coded into a computer. EMMALAB allows the user to progress completely through the mixing model and return to previous stages of the process to determine the ideal set of tracers and endmembers to represent the stream chemistry data.

This section walks through the steps in the process of EMMA, using the Provo River as an example dataset, to set up and demonstrate the general workflow of EMMALAB, and illustrate the benefits of mixing in tracer space by comparing the tracer space model results with three different PCA-based models that follow the traditional EMMA method.

3.1 Introduction to EMMALAB and suggested workflow

EMMA is perhaps the most sophisticated hydrograph separation method available, ostensibly capable of accommodating simple to relatively complex watersheds. However, our impression is that EMMA is not as widely employed as it probably should be, because there is no standard, user-friendly software for implementing these models. Therefore, most EMMA users have implemented custom calculations in MATLAB, R, Python, or similar programming environments (Chaves et al., 2008; He et al., 2020; Petermann et al., 2018). These types of custom software routines can make the necessary model iterations for selecting tracers and endmembers inconvenient. To address this problem, we created EMMALAB to streamline the EMMA process and update the approach to mixing calculations. We developed the software application in MATLAB, with standalone versions available that do not require a MATLAB license to run. When using the app, the workflow is well defined to avoid

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

user error and create a more uniform approach to the calculation. The program and supporting materials are provided in Thompson et al. (2025).

EMMALAB streamlines our recommended workflow to allow quick implementation of trial-and-error modelling decisions and error minimization. The following is a typical workflow.

1. Organize stream and endmember chemistry samples for the area of interest. Arrange the data according to the Microsoft Excel template included with the app supporting materials, and load the data into the app. The current version of EMMALAB does not have the ability to estimate unknown endmember compositions, so the user must have measured potential endmembers to use the app (Xu Fei and Harman, 2022).

2. Categorize stream location(s) and endmembers with a unique symbol and colour for use in the PCA and other plots within the app.

3. Determine which isotopes and solutes in the dataset should be used in the initial tracer set, starting with a list of isotopes and solutes that are expected to behave conservatively. Be aware of the fundamentally different processes affecting isotopes and solutes to determine whether isotopes should be included in the analysis. This approach is streamlined within the app, allowing the user to view bivariate plots and associated correlation coefficients (R) between the pairs of isotopes and solutes, following the methods in Hooper (2003).

4. Perform a PCA in the app to view the stream and endmember data in U-space. Determine the number of PCs to retain by considering eigenvalues and the percent variance explained by each PC (Sections 2.7-2.8). Although the PC scores are not used if the mixing calculations are done in tracer space, it is still important to evaluate the PCA results to select endmembers.

5. Select a set of endmembers that (at least approximately) circumscribe the stream data in U-space. After the endmembers are selected, the mixing calculation may be performed.

6. Perform the mixing calculation by using a non-negative least-squares optimization using z-scored tracer values for both the stream and endmember data. EMMALAB automatically z-scores selected stream data and tracer values. Ideally, the system of equations should be either critically determined or overdetermined. The option to calculate in U-space using 2 PC or 3 PC space is available in the app but is not recommended. It is, however, demonstrated below in a model comparison (section 3.4).

7. Assess how well the model predicts the measured tracer values in the stream data. Rerun the model with a variety of potential conservative tracers not included in the initial selection and assess the model predictions to determine whether such tracers should be included in the final model. Systematically underpredicted or overpredicted values may indicate nonconservative tracer behaviour, a poorly characterized endmember, or a missing endmember.

8. Once the final tracer set is selected, reconsider all measured endmembers. In some cases, a potential endmember could have been initially excluded because it appears in U-space to be a mixture of two other endmembers. However, in higher-dimensional space it may not actually lie between the two other endmembers and so including it could contribute to better model predictions. It may also be that there is some a priori reason (e.g., knowledge about the local geology) to include the initially excluded endmember. Therefore, the analyst can experiment with modelling the system using expanded sets of endmembers, iteratively evaluating the quality of the model predictions.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

3.2 Conservative tracer selection

As mentioned previously, it can be important to select as many tracers as possible, because this mathematically allows for the inclusion of more endmembers. In addition, constraining the model with more tracers can highlight possible systematic errors in the endmember characterizations (see Section 3.5 below). However, there has traditionally been minimal guidance regarding the process for selecting conservative tracers for EMMA. For instance, Christophersen et al. (1990a) suggested simply selecting the tracers that usually behave conservatively in streams, whereas Hooper (2003) suggested creating bivariate plots and choosing a set of tracers that exhibit a high degree of linear correlation with one another. The latter method has been critiqued as potentially misleading because linear or nonlinear trends do not necessarily guarantee conservative or nonconservative behaviour, respectively (James and Roulet, 2006; Barthold et al., 2011). For instance, weak linearity in a bivariate plot might simply indicate that the supply of the two potential tracers in question is dominated by more than two endmembers. Nevertheless, while high correlation between two solutes is not a guarantee that they behave conservatively, the probability that high linear correlation between several tracers is a simple fluke becomes progressively lower as more tracers are added. Therefore, linear correlation at least seems like a good place to start.

We recommend following an iterative procedure for selecting tracers in EMMALAB that begins with examining the correlation between potential tracers. Bivariate plots of the elements and/or isotopes should be created to assess the degree of linear correlation between potential tracers. Although there is no a priori method for determining an acceptable degree of correlation, Hooper (2003) suggests $R > 0.50$. An arbitrarily set R value may be used for selecting a group of tracers as long as it is applied consistently in the initial set of solutes and isotopes. In our case, we applied a more stringent criterion of $R > 0.80$ because of high linearity between several variables in our stream chemistry data. Oxygen and hydrogen isotopes should be treated as a separate category because their variation is governed by different processes than solutes. One strategy would be to select $\delta^{18}O$ and $\delta^{2}H$ unless they are poorly correlated, and then select other common tracers, e.g., Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Si, Cl^{-}, HCO_3^{-}, and electrical conductivity (Barthold et al., 2011). Within this initial group of non-isotopic tracers, those with low R value relative to the others should initially be excluded to retain a set that are all highly correlated with one another. Excluded tracers can subsequently be revisited later in the process. Although this method likely will not identify all conservative tracers in the dataset, it allows a starting point to establish an initial set of tracers that are most likely to behave conservatively.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

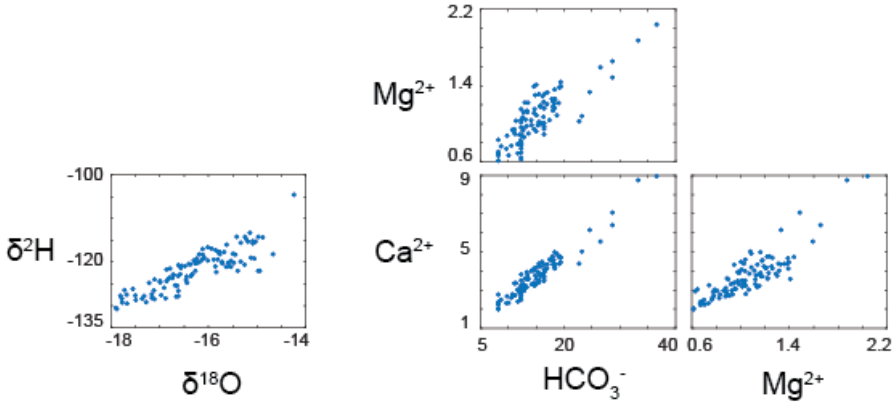


Figure 3: Bivariate plots of the tracers used in the initial step of tracer selection in EMMA for Provo River. Each solute displayed plausibly conservative behaviour (R > 0.80). $\delta^{18}O$ and $\delta^{2}H$ did not exhibit linear relationships with the solute tracers since fundamentally different processes explain variations in solute concentrations and isotopic ratios. These solutes are used to establish a foundation for the model which may be built upon in the second step of tracer selection by iteratively testing additional tracers through the model. Units for each solute are mg/L, and units for the isotopic ratios are permil.

The second phase of tracer selection is undertaken later in the process, after an initial mixing model has been calculated. The analyst optimizes the model with this initial set of tracers, then reoptimizes after iteratively including more potential tracers, one at a time. When a new potential tracer is added and the model reoptimized, the analyst may retain it for the next iteration if the model predicts the new tracer approximately as well or better than the initial set. Likewise, if at any point in this process adding more potential tracers results in good prediction of most tracers, but poorer prediction of one in the initial set, that previously selected tracer can be removed. This allows the analyst to tease out any remaining conservative tracers despite a lower degree of linear correlation in the bivariate plots. This procedure may be repeated at various points in the modelling process, and the quality of results of each iteration can be compared via the correlation between predicted and measured concentrations (R^2), deviation of the slope from 1, the root mean squared error (RMSE) of the predicted tracer concentrations, and by looking for systematic deviations in the residuals. Applying this procedure to the Provo River data, we first selected $\delta^{18}O$, $\delta^{2}H$, HCO_3^-, Mg^{2+}, and Ca^{2+} because the isotopes were well correlated with each other, as were the solutes (Fig. 3). We followed the subsequent steps of the modelling procedure (see below) using this initial set of tracers, which were all predicted well by the mixing model. We then iteratively included more potential tracers, and although they were not well correlated with the initial set of solute tracers, we found K^+, Si, Cl^-, and Na^+ were worth considering. When included, K^+ was extremely well predicted, and so was Si in most stream samples. However, in several samples Si was systematically over- or under-predicted during certain seasons (see Section 3.5). Na^+ and Cl^- exhibited poor correlation with the initial set of tracers but were moderately well predicted when included. They were ultimately excluded, however, because including them decreased model performance predicting Si and K^+, and we had reason to believe that road salt contamination was occurring (see Section 3.5). (We did not know how to characterize a hypothetical road salt endmember.) This tracer re-evaluation process occurred at

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

several points in the analysis and was invaluable, for example, for deciding whether an adequate number of endmembers had been chosen, or whether the chosen endmembers were adequately characterized.

These results illustrate certain points about the difficulty involved in choosing a maximal, or even adequate, set of tracers. First, the fact that Na^+ and Cl^- were ultimately rejected due in part to the likelihood of road salt contamination shows that the strategy employed in some studies to choose tracers simply because they typically behave conservatively, or even because they have been shown to behave conservatively upstream in the same watershed (Guinn et al., 2010; Burns et al., 2001), should be used with caution. Second, iteratively including initially rejected tracers can be important not only for expanding potential model complexity, but also to sufficiently constrain the model to highlight systematic deviations from model predictions, such as those revealed for Si (see Section 3.5). In contrast, some EMMA analysts have rejected tracers *because* they were highly correlated with a retained tracer and therefore deemed unnecessary (Lund et al., 2023). Finally, revisiting tracer selection later in the modelling process can be important for re-evaluating chosen endmembers as a maximal set of tracers is sought. However, some analysts have instead reduced the number of potential tracers to allow the previously chosen endmembers to be more adequately modelled (Liu et al., 2004; Montagud et al., 2021). This approach may or may not be justified in any particular case but it is worth pointing out that endmember selection and/or description is at least as likely a culprit for inadequate model results. All the above points illustrate the need for EMMA software capable of accommodating rapid model iteration.

3.3 Endmember selection

EMMA abstracts and reduces the diverse processes occurring in a watershed into a simple mixing problem involving a few endmember waters. And while tracer selection is important, one can usually find multiple solutes and isotopes that behave essentially conservatively in a system of interest. Therefore, the success of an EMMA model largely depends upon the proper selection and description of endmembers.

The first step in endmember selection is to perform PCA on the stream chemistry data, transform the endmember data into the same coordinate system, and plot both the endmember and stream data in the reduced dimensional space (U-space). The purpose of this is to choose potential endmembers that circumscribe (at least most of) the mixture data (Christophersen and Hooper, 1992), when plotted against combinations of the retained axes (PCs). In the case of the Provo River dataset, we decided to retain 3 PCs.

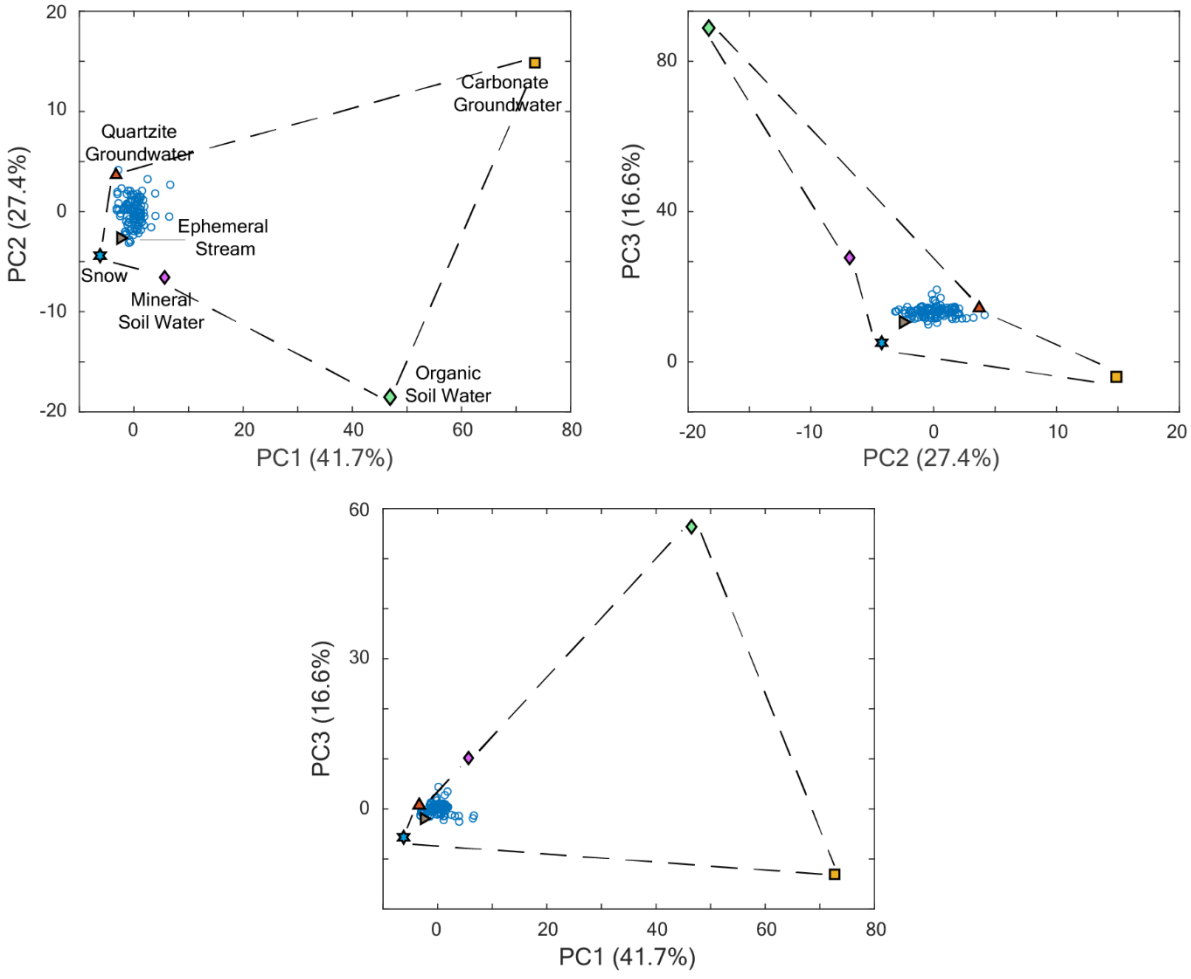
Why did we choose to retain three PCs? The methods typically used to decide how many PCs should be retained (see Sections 2.7-2.8) can give contradictory results; for example, the rule of one (retaining PCs with eigenvalues > 1) supported retaining three PCs, whereas a scree test (retaining PCs up to an "elbow" in the cumulative variance explained, see Fig. S1) supported retaining four PCs. However, since the first three PCs explained 85.6% of the variance in the original, *z*-scored stream dataset, and 3 PCs can be visualized simultaneously, we subjectively chose to retain only those (Fig. 4). It should also be mentioned that when the mixing calculation is performed in tracer space rather than U-space, the choice of the exact number of PCs to retain only matters for reconnaissance for potential endmembers. We found that quartzite groundwater, carbonate groundwater, snow, mineral soil water, and organic soil water sufficiently circumscribed the river data. Although the data could be

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

circumscribed without the inclusion of organic soil water, we decided to retain this endmember since it was uniquely enriched in K^+.

The mixing calculations (see Sections 2.9 and 3.4) pose some constraints on the number of endmembers that can be selected. The minimum number of endmembers needed is one more than the number of PCs retained, because this is the number of points required to form a convex hull around the stream data, and it also results in a critically determined mixing calculation (Hooper, 2003; Barthold et al., 2010; Foks et al., 2018; Pelizardi et al., 2017). Likewise, the maximum number of endmembers is also one more than the number of PCs retained if the mixing calculation is done in U-space. However, it is one more than the number of tracers (which is always more than the number of retained PCs) if the mixing calculation is done in tracer space. Using a smaller number of endmembers than is possible when mixing in tracer space results in an overdetermined system of equations, for which a best-fit solution can be found using non-negative least-squares optimization (Gentle, 1998).

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025



515 **Figure 4: Principal components analysis for the Provo River data from Soapstone (blue circles) with potential endmembers projected into the river-defined U-space. The endmembers that adequately circumscribe the data are quartzite groundwater, snow, mineral soil water, organic soil water, and carbonate groundwater. The plots show PC1 versus PC2, PC2 versus PC3, and PC1 versus PC3, which altogether explain 85.7% of the variance within the dataset.**

520 **3.4 Mixing calculation**

Here we compare the results of four mixing calculations using the five endmembers we selected for the Provo River (see Section 3.3) to show that constraining the system with more variables (e.g., calculating in the full tracer space) generally results in more robust optimizations. The first model was calculated using our seven selected tracers (see Section 3.2), while the calculation for the three other models was done in U-space with two (69.1% variance), three (85.7%), and four (96.4%) PCs.

525 In Fig. 5 we show the difference between the calculated fractional contributions of all endmembers in the U-space mixing models from those calculated from the tracer-space model. As we include more PCs (and hence more of the information from

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

the original data), we get consistently closer to the results of the tracer-space model. The maximum differences in the fractional contributions of all endmembers between the 2 PC, 3 PC, and 4 PC models were (-0.28, 0.24), (-0.11, 0.14), and (-0.03, 0.03), respectively, with mineral soil water and snow being the most consistently over- and under-predicted endmembers (Fig. 5). The predicted vs. observed tracer plots for these models (see Figs. S2 – S4) show that some of the tracers become increasingly better predicted as the calculations are based on more information, with high R^2 and low y-intercept and slope-1 produced by the tracer space model.

Given these results, it is clear that performing the mixing calculation in tracer space reliably simplifies the approach to EMMA, while allowing for greater model complexity by constraining the model with more variables.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

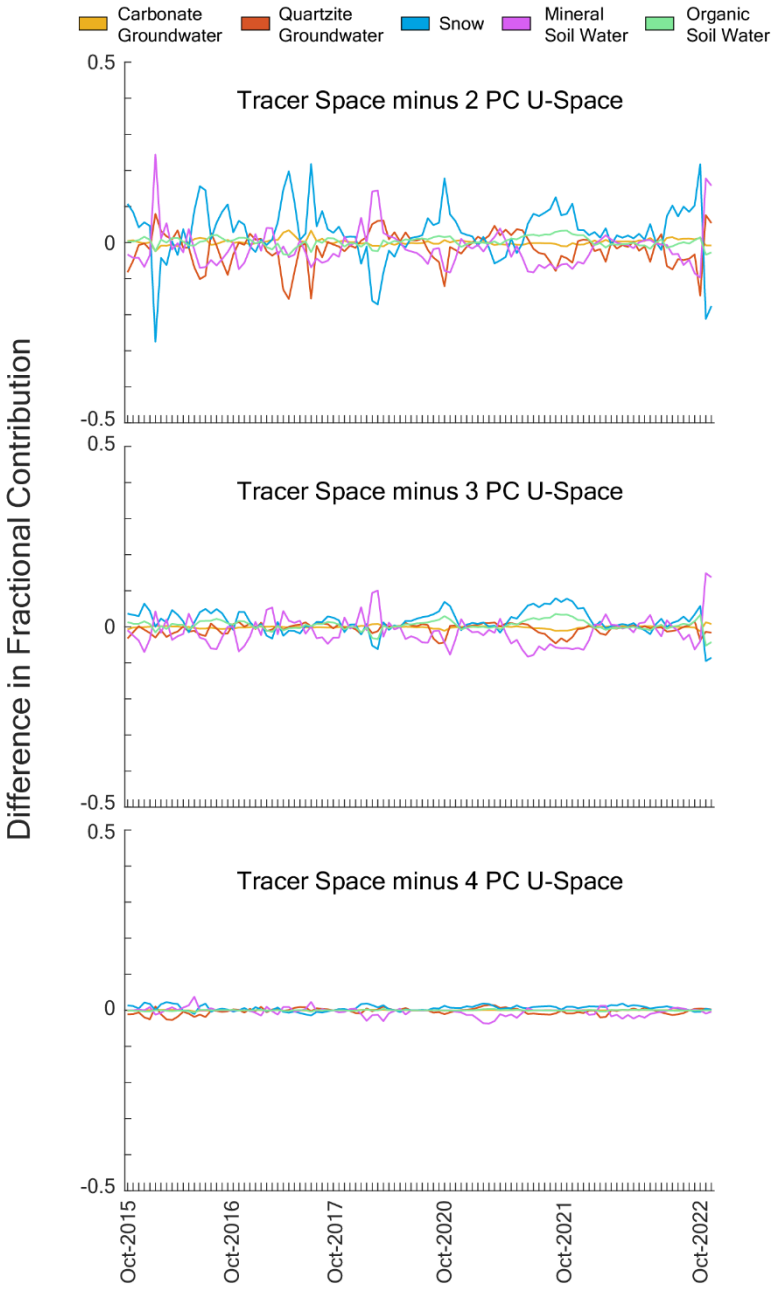


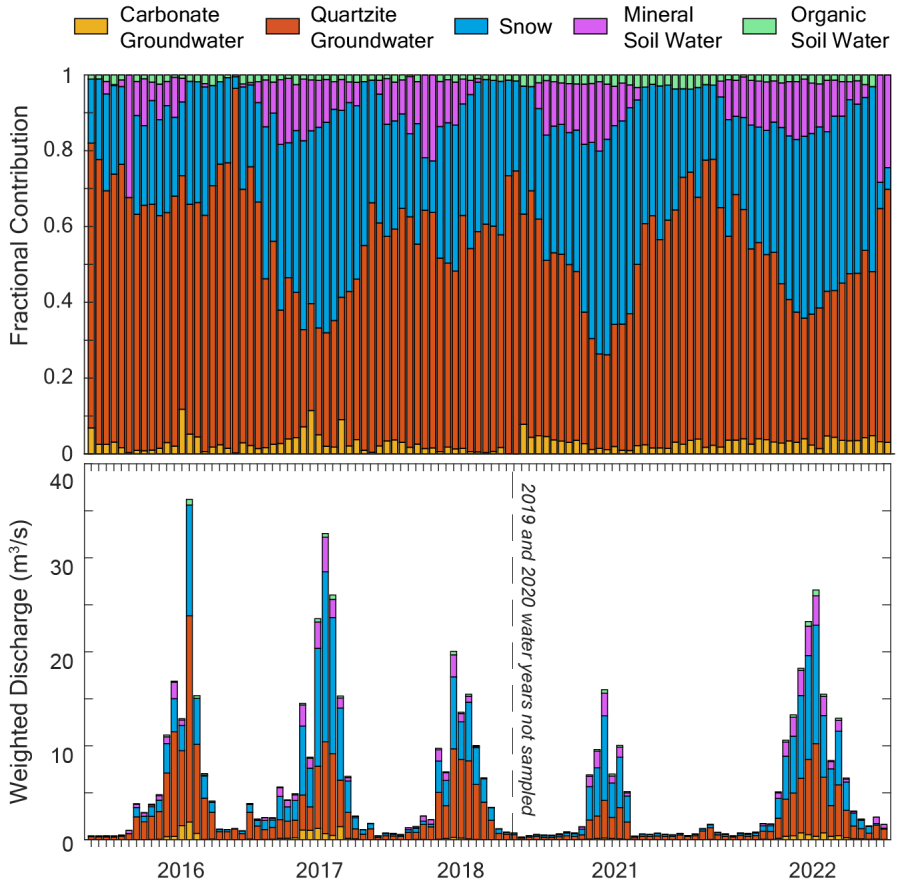
Figure 5: Plots of the fractional contributions calculated using 2 PCs, 3 PCs, 4 PCs, each subtracted by the fractional contributions calculated from the tracer space model.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

3.4.1 Model results

The 7-tracer mixing model was deemed most accurate for assessing the fractional contributions of each endmember to the Provo River. Using this model, the fractional endmember contributions at Soapstone showed strong seasonal variability in the fractional contributions of quartzite groundwater, snow, mineral soil water, organic soil water, and carbonate groundwater (Figure 7). During baseflow (December – February), the river was dominated by quartzite groundwater, associated with the most common lithology upstream of Soapstone. The snow endmember increased greatly during snowmelt runoff (April – June), with maximum contributions ranging between 22 – 49% across multiple years, with an average cumulative spring runoff snow input of 44.5 m^3/s. The increase in the snow fraction occurred at the same time as snowmelt runoff, but the total volume of snow input into the stream each year was difficult to quantify. The 2017 water year had above-average snowpack, which was reflected in increased calculated fractional contributions of snow, but similar trends were seen in 2021, which was a below-average snowpack year. Thus, high snowpack does not necessarily result in higher fractional snow input in the model, although the total discharge was clearly lower in 2021. This could occur because antecedent soil moisture and air temperature also control the amount of snowmelt making it directly into the stream without significant soil interaction (Brooks et al., 2021). Indeed, there was an increase in soil water contributions (both types) during snowmelt runoff, likely caused by snowmelt flushing the soils.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025



555 **Figure 6: Fractional contributions and weighted discharge from each endmember for Provo River for 2016 – 2018 and 2021 – 2022.**

3.5 Model tracer predictions

As suggested by Christophersen et al. (1990a), the accuracy of model-predicted tracer concentrations may be used to test the plausibility of the combination of endmembers included. A poor model prediction (low R^2, high slope-1 and y-intercept, or 560 systematic errors) may suggest inadequately characterized endmembers, non-conservative behaviour by some of the tracers, missing endmembers, or inadequate sampling of endmembers (Foks et al., 2018).

In the initial tracer selection process, Na^+ and Cl^- were considered for the model since these solutes typically behave conservatively, but they were systematically underpredicted, suggesting a missing endmember. However, there appeared to be no pattern in the underpredicted values. Along the highway upstream from Soapstone, the road is regularly maintained in early 565 spring with snowplows and road salt. This opens the possibility that the underpredicted values of Na^+ and Cl^- could be caused by the transfer of road salt into the river, likely through snowmelt. Given that this missing endmember comes from a sporadic source (road salt) not associated with any particular amount of water, the difficulty in estimating the input makes it an

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

unsuitable endmember for EMMA modelling. In the case of the Provo River mixing calculations, the final model-predicted concentrations of $\delta^{18}O$, $\delta^{2}H$, Si, HCO_3^-, Mg^{2+}, K^+, and Ca^{2+} were compared to the observed stream chemistry data to assess model accuracy (Figure 8). The R^2, slope-1, y-intercept, and RMSE were used to assess each plot and determine how well the model was able to replicate the stream data based on the selected endmembers (Table 2). We primarily considered the predictions to be good if the RMSE was relatively small. The statistics for the fit of the plots can be found in Table 2. Of the predictions, $\delta^{2}H$ and K^+ were predicted very well, $\delta^{18}O$, HCO_3^-, and Mg^{2+} were underpredicted, Ca^{2+} was overpredicted, and Si was both under- and over-predicted depending on time of year. In the case of Si, we observed a systematic deviation where it was overpredicted primarily during early baseflow, and occasionally underpredicted during peak runoff (Fig. S5). The overpredicted Si values during late summer may suggest uptake in soils during peak photosynthesis (Van Breemen et al., 2000). Sedges and moss, in particular, are efficient Si-accumulators (Ma and Takahashi, 2002; Hodson et al., 2005), and are found in the watershed. The underpredicted values during runoff may be caused by soil-water sampling having been conducted during mid-late runoff season (late May-June) and may not have captured maximum Si concentrations earlier in the year (Gérard et al., 2002).

Other systematic deviations of tracer values from the model predictions are also readily explainable. For example, while the $\delta^{2}H$ values are predicted very well, the $\delta^{18}O$ values in the river are systematically underpredicted, perhaps from evaporation. Systematically underpredicted HCO_3^- at high concentrations may indicate CO_2 degassing in the stream. The underpredicted Mg^{2+} and overpredicted Ca^{2+} concentrations may be related to variability in the Mg and Ca content of limestones and dolostones in the sedimentary bedrock, which would cause differences in these as groundwater interacts with different areas and depths of bedrock seasonally.

These examples illustrate an important point. That is, maximizing the number of tracers included and optimizing the mixing in tracer space leads to maximally constrained models. This can lead to less scatter in the predicted vs. observed tracer plots, but simultaneously can make systematic deviations more apparent. If the systematic deviations can be readily explained in physical terms, the value of the model often increases. For example, it is possible that our carbonate groundwater endmember samples were collected from locations where the water had interacted with more limestone, as opposed to dolostone, than would be representative of similar waters in the watershed. If so, then including Mg^{2+}, Ca^{2+}, and HCO_3^- as tracers likely resulted in better estimates of the carbonate groundwater contributions. Similarly, including Si did not make the predictions of the other tracers worse, and its systematic deviation from predicted concentrations could inform future soil water sampling. Various applications of Si and K^+ to the model illustrate that both tracers were essential in accurately defining both mineral and organic soil water (Fig. S6).

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

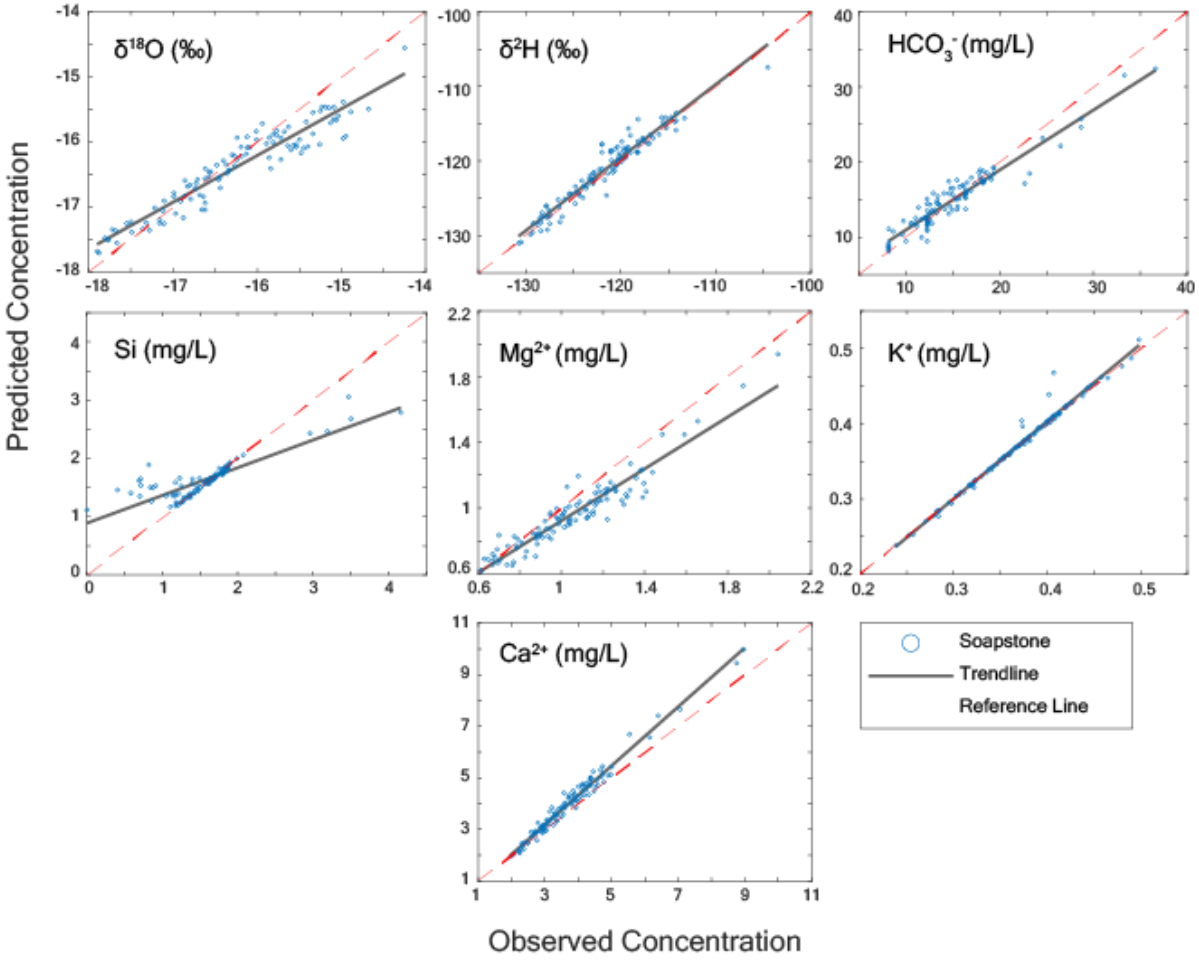


Figure 7: Model predicted concentrations versus observed concentrations for Provo River using five endmembers.

Table 2: Statistics associated with the fit between the model predicted concentrations and observed concentrations.

	R^2	Slope-1	Intercept	RMSE
d18O	0.9	-0.28	-4.72	0.34
d2H	0.93	-0.02	-2.33	1.36
HCO3	0.9	-0.2	2.98	1.64
Si	0.7	-0.53	0.92	0.36
Mg	0.9	-0.21	0.13	0.13
K	0.98	0.03	-0.01	0.01

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

Ca	0.98	0.14	-0.26	0.39

3.6 Error estimation

Jackknifing endmember means resulted in 17,820 different combinations of the means of the five endmembers, which we used to assess the uncertainty in the fractional endmember contributions to the individual stream samples (Fig. 9). The calculated errors for all fractional contributions were small. The mineral soil water endmember had the greatest amount of uncertainty (+/- 7%), and quartzite groundwater had the smallest (+/- 0.5%). This was expected because soil water included a small number of samples with large variability in water chemistry, which would lead to a larger variance in jackknifed means. Soil water endmembers can generally be difficult to adequately capture because soils can dramatically differ spatially, between soil types, and via different chemical and biological processes (Christophersen et al., 1990b).

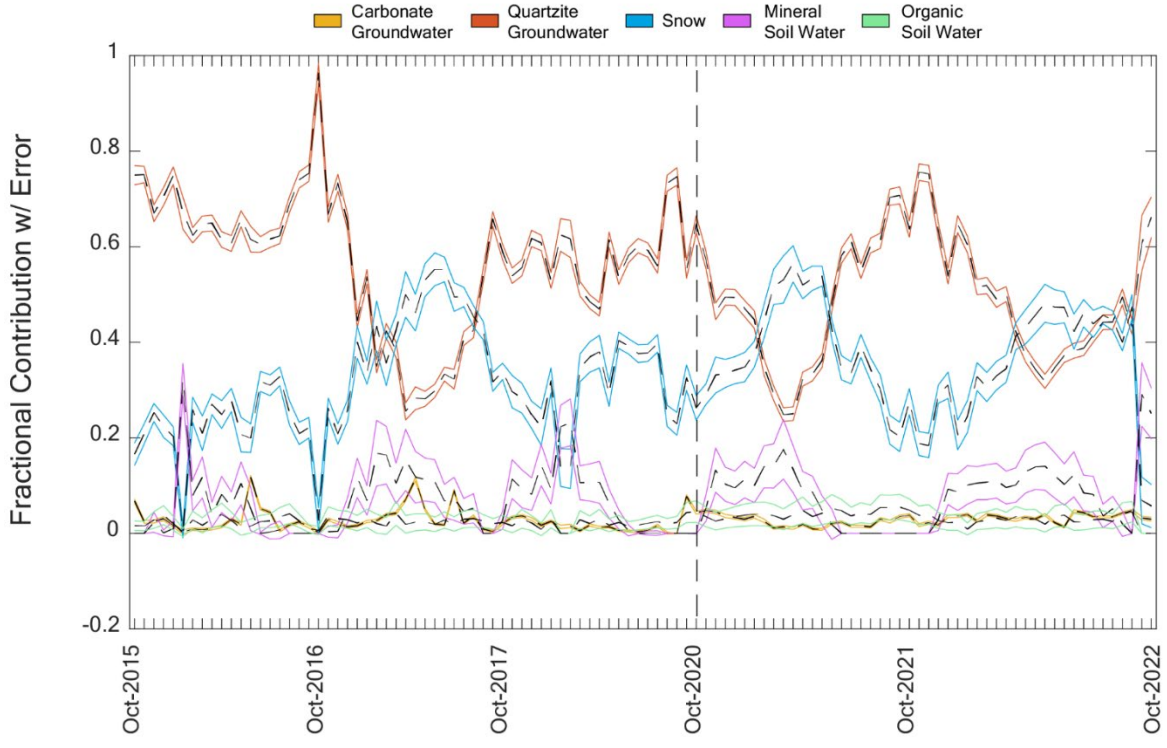


Figure 8: Error within endmember sampling groups for each sampling day for Soapstone, with shaded areas representing the extent of error within the endmembers. The dashed line indicates a break in the sampling in the 2019 and 2020 water years.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

4 Conclusions

Endmember mixing analysis is a sophisticated and useful tool for understanding how the primary water sources of a stream change over time. However, EMMA is underused by the hydrological community and has sometimes been improperly implemented because there has been no standard software for it. Here we have introduced an open-source software package, EMMALAB, which allows for rapid model setup and iteration, and have described a systematic, iterative approach to applying EMMA to take advantage of this capability. This approach includes a revised method for performing mixing calculations in tracer space, rather than reduced dimensional space, which allows for the inclusion of more potential water sources and better constrained models.

We illustrated our approach with data from the upper Provo River to demonstrate the utility of this approach. In this example, after several iterations we found that seven tracers and five endmembers were adequate to describe the water chemistry of our sampling location. These variables were applied to the watershed to estimate the relative contributions of each endmember over the five years of data collected. We found that snow was consistently the main contributor during spring runoff, with an average cumulative input during spring runoff of 44.5 m^3, with quartzite groundwater being the second highest contributor (43.8 m^3). Quartzite groundwater was the primary contributor during baseflow (0.65 m^3). This application illustrates the streamflow response to spring snowmelt and the importance of snow to the hydrologic processes within the watershed. A continued application of this analysis to Provo River could help to understand the importance regular snowpack deposition has on the health of the watershed.

Code/Data availability

All data and EMMALAB software application resources used in this study are freely available through HydroShare at https://www.hydroshare.org/resource/90ad78faec9f41c180d9057b9e815785/ (Thompson et al., 2025).

Author contribution

ANT, BRB and GTC wrote the manuscript. ANT, BRB and EJE designed and implemented the mathematical procedures in the software package. ANT and KLC collected field samples and performed the lab analyses. JJL, STN, KAR, and DPF assisted with data interpretation.

Competing interests

The authors declare that they have no conflict of interest.

Financial support

This work was supported by the U.S. National Science Foundation grants EAR-2012093 and EAR-2005432.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

References

Albarède, F.: Geochemistry: An Introduction, 2nd, Cambridge University Press, Cambridge, 26-28 pp., 10.1017/cbo9780511807435.006, 2009.

Ali, G. A., Roy, A. G., Turmel, M.-C., and Courchesne, F.: Source-to-stream connectivity assessment through end-member mixing analysis, Journal of Hydrology, 392, 119-135, 10.1016/j.jhydrol.2010.07.049, 2010.

Barthold, F. K., Tyralla, C., Schneider, K., Vaché, K. B., Frede, H.-G., and Breuer, L.: How many tracers do we need for end member mixing analysis (EMMA)? A sensitivity analysis, Water Resources Research, 47, 10.1029/2011wr010604, 2011.

Barthold, F. K., Wu, J., Vaché, K. B., Schneider, K., Frede, H.-G., and Breuer, L.: Identification of geographic runoff sources in a data sparse region: hydrological processes and the limitations of tracer-based approaches, Hydrological Processes, 24, 2313-2327, 10.1002/hyp.7678, 2010.

Bazemore, D. E., Eshleman, K. N., and Hollenbeck, K., J.: The role of soil water in stormflow generation in a forested headwater catchment: synthesis of natural tracer and hydrometric evidence, 1994.

Birkel, C., Correa Barahona, A., Duvert, C., Granados Bolaños, S., Chavarría Palma, A., Durán Quesada, A. M., Sánchez Murillo, R., and Biester, H.: End member and Bayesian mixing models consistently indicate near-surface flowpath dominance in a pristine humid tropical rainforest, Hydrological Processes, 35, 10.1002/hyp.14153, 2021.

Brooks, P. D., Gelderloos, A., Wolf, M. A., Jamison, L. R., Strong, C., Solomon, D. K., Bowen, G. J., Burian, S., Tai, X., Arens, S., Briefer, L., Kirkham, T., and Stewart, J.: Groundwater-Mediated Memory of Past Climate Controls Water Yield in Snowmelt-Dominated Catchments, Water Resources Research, 57, 10.1029/2021wr030605, 2021.

Bryant, B.: Geologic map of the Salt Lake City 30' x 60' quadrangle, north-central Utah and Uinta County, Wyoming, Utah Geological Survey, 2003.

Bugaets, A., Gartsman, B., Gubareva, T., Lupakov, S., Kalugin, A., Shamov, V., and Gonchukov, L.: Comparing the Runoff Decompositions of Small Experimental Catchments: End-Member Mixing Analysis (EMMA) vs. Hydrological Modelling, Water, 15, 10.3390/w15040752, 2023.

Burns, D. A., McDonnell, J. J., Hooper, R. P., Peters, N. E., Freer, J. E., Kendall, C., and Beven, K.: Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research Watershed (Georgia, USA), Hydrological Processes, 15, 1903-1924, 10.1002/hyp.246, 2001.

Carrera, J., Vázquez-Suñé, E., Castillo, O., and Sánchez-Vila, X.: A methodology to compute mixing ratios with uncertain end-members, Water Resources Research, 40, 10.1029/2003wr002263, 2004.

Cattell, R. B.: The Scree Test For The Number Of Factors, Multivariate Behav Res, 1, 245-276, 10.1207/s15327906mbr0102_10, 1966.

Chaves, J., Neill, C., Germer, S., Neto, S. G., Krusche, A., and Elsenbeer, H.: Land management impacts on runoff sources in small Amazon watersheds, Hydrological Processes, 22, 1766-1775, 10.1002/hyp.6803, 2008.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

675 Checketts, H. N., Carling, G. T., Fernandez, D. P., Nelson, S. T., Rey, K. A., Tingey, D. G., Hale, C. A., Packer, B. N., Cordner, C. P., Dastrup, D. B., and Aanderud, Z. T.: Trace Element Export From the Critical Zone Triggered by Snowmelt Runoff in a Montane Watershed, Provo River, Utah, USA, Frontiers in Water, 2, 10.3389/frwa.2020.578677, 2020.

Chisola, M. N., van der Laan, M., and Butler, M. J.: Quantifying streamflow sources to improve water allocation management in a catchment undergoing agricultural intensification, Physics and Chemistry of the Earth, Parts A/B/C, 128,
680 10.1016/j.pce.2022.103227, 2022.

Christophersen, N. and Hooper, R. P.: Multivariate analysis of stream water chemical data: The use of principal components analysis for the end-member mixing problem, Water Resources Research, 28, 99-107, 10.1029/91wr02518, 1992.

Christophersen, N., Neal, C., Hooper, R. P., Vogt, R. D., and Andersen, S.: Modelling streamwater chemistry as a mixture of soilwater end-members - An application to the Panola Mountain catchment, Georgia U.S.A., Journal of Hydrology, 116,
685 321-343, 1990a.

Christophersen, N., Neal, C., Hooper, R. P., Vogt, R. D., and Anderson, S.: Modelling streamwater chemistry as a mixture of soilwater end-members - A step towards second-generation acidification models, Journal of Hydrology, 307-320, 1990b.

Clark, I. D. and Fritz, P.: Environmental Isotopes in Hydrogeology, CRC Press/Lewis Publishers, Boca Raton, FL1997.

Cuoco, E., Viaroli, S., Darrah, T. H., Paolucci, V., Mazza, R., and Tedesco, D.: A geometrical method for quantifying
690 endmembers' fractions in three-component groundwater mixing, Hydrological Processes, 35, 10.1002/hyp.14409, 2021.

Davis, J. C.: Statistics and Data Analysis in Geology, 3rd, John Wiley & Sons2002.

Durand, P. and Torres, J. L. J.: Solute transfer in agricultural catchments: the interest and limits of mixing models, 1996.

Foks, S. S., Stets, E. G., Singha, K., and Clow, D. W.: Influence of climate on alpine stream chemistry and water sources, Hydrological Processes, 32, 1993-2008, 10.1002/hyp.13124, 2018.

695 Genereux, D.: Quantifying uncertainty in tracer-based hydrograph separations, Water Resources Research, 34, 915-919, 10.1029/98wr00010, 1998.

Gentle, J. E.: Numerical linear algebra for applications in statistics, Springer-Verlag New York, Inc.1998.

Gérard, F., François, M., and Ranger, J.: Processes controlling silica concentration in leaching and capillary soil solutions of an acidic brown forest soil (Rhône, France), Geoderma, 107, 197-226, 10.1016/s0016-7061(01)00149-5, 2002.

700 Guinn, C. G., Vulava, V. M., Callahan, T. J., Jones, M. L., and Ginn, C. L.: Using water chemistry data to quantify source contribution to stream flow in a coastal plain watershed, 2010.

He, Z., Unger-Shayesteh, K., Vorogushyn, S., Weise, S. M., Duethmann, D., Kalashnikova, O., Gafurov, A., and Merz, B.: Comparing Bayesian and traditional end-member mixing approaches for hydrograph separation in a glacierized basin, Hydrology and Earth System Sciences, 24, 3289-3309, 10.5194/hess-24-3289-2020, 2020.

705 Hodson, M. J., White, P. J., Mead, A., and Broadley, M. R.: Phylogenetic variation in the silicon composition of plants, Ann Bot, 96, 1027-1046, 10.1093/aob/mci255, 2005.

Hooper, R. P.: Diagnostic tools for mixing models of stream water chemistry, Water Resources Research, 39, 10.1029/2002wr001528, 2003.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

Jacobs, S. R., Timbe, E., Weeser, B., Rufino, M. C., Butterbach-Bahl, K., and Breuer, L.: Assessment of hydrological pathways in East African montane catchments under different land use, Hydrology and Earth System Sciences, 22, 4981-5000, 10.5194/hess-22-4981-2018, 2018.

James, A. L. and Roulet, N. T.: Investigating the applicability of end-member mixing analysis (EMMA) across scale: A study of eight small, nested catchments in a temperate forested watershed, Water Resources Research, 42, 10.1029/2005wr004419, 2006.

Jolliffe, I. T. and Cadima, J.: Principal component analysis: a review and recent developments, Philos Trans A Math Phys Eng Sci, 374, 20150202, 10.1098/rsta.2015.0202, 2016.

Joreskog, K. G., Klovan, J. E., and Reyment, R. A.: Geological Factor Analysis, Elsevier Scientific Publishing Company1976.

Lichtert, S. and Verbeeck, J.: Statistical consequences of applying a PCA noise filter on EELS spectrum images, Ultramicroscopy, 125, 35-42, 10.1016/j.ultramic.2012.10.001, 2013.

Liu, F., Williams, M. W., and Caine, N.: Source waters and flow paths in an alpine catchment, Colorado Front Range, United States, Water Resources Research, 40, 10.1029/2004wr003076, 2004.

Liu, G., Ma, F., Liu, G., Guo, J., Duan, X., and Gu, H.: Quantification of Water Sources in a Coastal Gold Mine through an End-Member Mixing Analysis Combining Multivariate Statistical Methods, Water, 12, 10.3390/w12020580, 2020.

Lukens, E., Neilson, B. T., Williams, K. H., and Brahney, J.: Evaluation of hydrograph separation techniques with uncertain end-member composition, Hydrological Processes, 36, 10.1002/hyp.14693, 2022.

Lund, J., Forster, R. R., Jameel, Y., Rupper, S. B., Deeb, E. J., Dars, G. H., Zaheer, A., Ali, M., Ghafoor, A., Khan, G., Arfan, M., Liston, G. E., Akhter Qureshi, J., Carling, G., and Burian, S. J.: Constraining Mountain Streamflow Constituents by Integrating Citizen Scientist Acquired Geochemical Samples and Sentinel-1 SAR Wet Snow Time-Series for the Shimshal Catchment in the Karakoram Mountains of Pakistan, Water Resources Research, 59, 10.1029/2022wr032171, 2023.

Ma, J. F. and Takahashi, E.: Chapter 5 - Silicon-accumulating plants in the plant kingdom, Soil, Fertilizer, and Plant Silicon Research in Japan, 63-71, https://doi.org/10.1016/B978-044451166-9/50005-1, 2002.

Miesch, A. T.: Scaling variables and interpretation of eigenvalues in principal component analysis of geologic data, Journal of the International Association for Mathematical Geology, 12, 523-538, 1980.

Montagud, D., Sala, M., and Camarero, L.: Applicability of mixing modelling to determine the contributions to surface flow in high mountain catchments, Hydrological Sciences Journal, 66, 2382-2394, 10.1080/02626667.2021.1985723, 2021.

Munroe, J. S., Attwood, E. C., O'Keefe, S. S., and Quackenbush, P. J. M.: Eolian deposition in the alpine zone of the Uinta Mountains, Utah, USA, Catena, 124, 119-129, 2015.

Munroe, J. S., Norris, E. D., Olson, P. M., Ryan, P. C., Tappa, M. J., and Beard, B. L.: Quantifying the contribution of dust to alpine soils in the periglacial zone of the Uinta Mountains, Utah, USA, Geoderma, 378, 10.1016/j.geoderma.2020.114631, 2020.

https://doi.org/10.5194/egusphere-2025-2053
Preprint. Discussion started: 6 June 2025

Pelizardi, F., Bea, S. A., Carrera, J., and Vives, L.: Identifying geochemical processes using End Member Mixing Analysis to decouple chemical components for mixing ratio calculations, Journal of Hydrology, 550, 144-156, 10.1016/j.jhydrol.2017.04.010, 2017.

Petermann, E., Knöller, K., Rocha, C., Scholten, J., Stollberg, R., Weiß, H., and Schubert, M.: Coupling End-Member Mixing Analysis and Isotope Mass Balancing (222-Rn) for Differentiation of Fresh and Recirculated Submarine Groundwater Discharge Into Knysna Estuary, South Africa, Journal of Geophysical Research: Oceans, 123, 952-970, 10.1002/2017jc013008, 2018.

Reclamation, U. S. B. o.: Provo River Project : Utah, Salt Lake, Summit, Utah, and Wasatch Counties, 1958.

Thompson, A. N., Bickmore, B. R., Evans, E. J., and Carling, G. T.: EMMALAB v. 1.1: Software for improved endmember mixing analysis [code], https://www.hydroshare.org/resource/90ad78faec9f41c180d9057b9e815785/, 2025.

Trauth, M. H.: MATLAB recipes for earth sciences, 5, Springer Nature Switzerland AG2021.

Valder, J. F., Long, A. J., Davis, A. D., and Kenner, S. J.: Multivariate statistical approach to estimate mixing proportions for unknown end members, Journal of Hydrology, 460-461, 65-76, 10.1016/j.jhydrol.2012.06.037, 2012.

van Breemen, N., Lunström, U. S., and Jongmans, A. G.: Do plants drive podzolization via rock-eating mycorrhizal fungi?, Geoderma, 94, 163-171, https://doi.org/10.1016/S0016-7061(99)00050-6, 2000.

Vonk, J. E., Sánchez-García, L., Semiletov, I. P., Dudarev, O. V., Eglinton, T. I., Andersson, A., and Gustafsson, Ö.: Molecular and radiocarbon constraints on sources and degradation of terrestrial organic carbon along the Kolyma paleoriver transect, East Siberian Sea, 10.5194/bgd-7-5191-2010, 2010.

Wold, S., Esbensen, K., and Geladi, P.: Principal component analysis, Chemometrics and Intelligent Laboratory Systems, 2, 37-52, https://doi.org/10.1016/0169-7439(87)80084-9, 1987.

Xu Fei, E. and Harman, C. J.: A data-driven method for estimating the composition of end-members from stream water chemistry time series, Hydrology and Earth System Sciences, 26, 1977-1991, 10.5194/hess-26-1977-2022, 2022.