A data-driven method for estimating the composition of end-members from stream water chemistry timeseries

by Esther Xu Fei, Ciaran J. Harman

Response to reviewers

September 20th 2021

We would like to thank the editor and reviewer for their thoughtful reviews, and for their patience as we have revised this paper. Below we respond to the individual concerns expressed by the reviewer

Major Concerns

Conservative behavior of tracers used in the analysis was not tested. I had this concern before, but it was not addressed adequately or correctly in the current revision. I suggested to run DTMM first to determine the conservative behavior of all solutes (e.g., completely conservative, semi- or quasi-conservative, and non-conservative under a lower dimensionality) and then to include conservative ones in CHEMMA. Instead, authors ran DTMM after CHEMMA and just compared the outcomes. They found that the residuals of sulfate, magnesium, and calcium still maintain some structures in a two-dimensional mixing space and the residual structures persist until the dimension goes beyond five (Figure 4 b). If this result is true (I am not sure if Shapiro-Wilk test is appropriate for this analysis; see my comment on Figure 4 below), it strongly indicates some solutes are not completely conservative. With six solutes, you cannot go beyond five dimensions to determine conservative solutes and the number of end-members (see my comment in the second round of revision). Authors misunderstood how DTMM works.

Thank you for these suggestions. It is important to note that the data used in the paper is a subset of precisely the same data (Panola Lower Gauge) as was used in the original DTMM paper (Hooper 2003). In other words, DTMM has already been applied to this dataset and the results published. Figure 7 in Hooper (2003) presents the concentration relation in the observational space, and Figure 9 in Hooper (2003) presents the residuals against concentration in one-, two-, and three-dimensional mixing space. Hooper (2003) does not conclude any solutes should be excluded from the analysis due to non-conservative mixing. Instead, he suggests four end-members (three-dimensional mixing space) may be warranted, which is consistent with the results presented here.

Given that our results are entirely consistent with the paper that first presented DTMM, and make use of the same dataset, we are unsure how to satisfy the concern of the reviewer. This issue also seems somewhat peripheral to the main contribution of the paper, which is to present the CHEMMA method. Certainly the choice of an appropriate number of dimensions is important, but it is a challenge for EMMA, CHEMMA, and all data

analytic techniques that rely on low-dimensional embeddings of high dimensional data. It is not the intention of this paper to solve that particular problem, and it seems unreasonable that the paper might not be accepted because it fails to due so.

With that said, I do not mean you cannot run CHEMMA with all six solutes together or with different groups of solutes. Instead, I do encourage authors to run different versions of CHEMMA with various combinations of solutes (following the DTMM results) and compare the outcomes, including the number of end-members.

While we appreciate the suggestion, we do not believe extending the length of the paper in this way would add anything of value. As we discuss above, there is no reason to exclude any of the solutes and so we do not see the point of running different combinations of them. The goal of this paper is to present the CHEMMA method, not to examine the influence of using different set of solutes.

As a research article, I strongly suggest to run DTMM first (following Hooper 2003), then EMMA (similar to Christophersen and Hooper 1992), and finally compare with the results of CHEMMA (e.g., groups of all six solutes and conservative solutes identified by DTMM as mentioned above). The comparison should not be limited to the end-member composition, but include the number of end-members, the fractional contributions of end-members, and the end-member distances. For example, how do the end-member distances from the end-member composition of CHEMMA compare to that of the measured ones? Is there an improvement in the end-member distances with CHEMMA? I did not keep track of whether or not Hooper (2003) used exactly the same data set as Hooper (1990). If so, you may not need to re-run both DTMM and EMMA but just summarize their results.

We thank the reviewer for these detailed suggestions. Many of these are already addressed in Christophersen and Hooper (1992), Hooper (2003), and in the present paper. End member distances for both EMMA and CHEMMA are shown in figure 3. The change in the fractional contributions of each end member can also be readily inferred from this figure (e.g. the fraction of water supplied by the CHEMMA end-member most like the organic end-member from Christophersen and Hooper (1992) is less than that inferred in the original study.

Beynd that, it is unclear to us what further application of EMMA and DTMM to this dataset will produce – certainly we cannot see how it better achieves the aims of the paper.

Addition of a test with varying sample sizes (Figure 6) is nice and very much appreciated. One result is very much promising (e.g., relatively stable compositions for end-members 2 and 3), but others are not (e.g., significant variation of composition of end-member 1; still many outliers for end-members 2 and 3). Together with significant variability in identifying end-members of the synthetic data (Figure 8) and high uncertainties of algorithm (Figure 7), it indicates that the data

structure or the distribution of sample points determines the end-member composition. The role of CHEMMA in end-member mixing analysis is limited. This limitation should be explicitly discussed and stated in the abstract and conclusion. This does not downplay CHEMMA's value, but simply tell the truth so that future users will not be misled. As a matter of fact, CHEMMA would be very helpful in identifying a missing end-member, guiding field sampling of end-members, and generating a hypothesis test.

We appreciate your sincere suggestion, and we have added clarifications in both abstract and conclusion. In Abstract, we changed the part after Line 12 to:

"We examine uncertainties in end-member identification arising from the random initialization of the CH-NMF algorithm, from the sample size, and from the data structure using both real and synthetic data. The results suggest that the robustness of the CHEMMA method depends on the dataset including, for each end-member, a subset of samples in which is nearly absent, as well as others in which it is a more substantial part."

We also added to the Conclusion at Line 355:

"However, the usefulness of CHEMMA is limited by the structure of the data in mixing space. As Figure 8 suggests, CHEMMA will fail for datasets in which all end members are present in all samples to some non-trivial degree. Samples in which an end-member is absent provide critical information, and strongly control the location of the face of the convex hull used to identify the other end members.

Moderate Comments/Concerns

Abstract:

The manuscript has been modified, with a new section (re: synthetic dataset) and a new analysis (re: varying number of samples) added, but the abstract was not updated to include the results from these analyses, nor was the limitation of the approach stated in the abstract.

Thank you. We have updated the abstract.

Introduction:

Somewhere in the introduction, all of the mixing model assumptions needs to be explicitly listed (i.e., i. Tracers used in the mixing model must be conservative; ii. The number of end-members is known; iii. End-member compositions must be distinct for at least one tracer; iv. End-member compositions are spatiotemporally constant or their variations are known or treated as different end-members). These assumptions should be discussed, e.g., which ones have been addressed by which tools and which ones are still up for research. In my opinion, the first two assumptions have been resolved by the diagnostic tools of mixing models, which has to be acknowledged to respect the earlier study. This will pave a clear pathway for your own research. However, this does not mean DTMM cannot be challenged or improved.

Thank you for your suggestions, we have added these assumptions to the end of the second paragraph in the Introduction.

"The EMMA method assumes that 1) solutes used in the mixing model are approximately conservative, 2) stream water consists of identifiable number of end-member sources, 3) end-member compositions are distinct for at least one tracer, and 4) end-member compositions are spatiotemporally constant (or their variations are known or can be approximated using additional end-members)."

Saying that including non-conservative solutes in the mixing models has not been resolved is inappropriate and misleading, as non-conservative solutes should not be included in the analysis based on the mixing model assumptions. This does not mean you cannot challenge the assumptions, but I do not think that is what your study aimed at.

Thank you for your suggestions, we have added a sentence after 3) in Line 73 that "therefore, only tracers that are believed to be approximately conservative (i.e. since they entered the stream their concentrations have been altered primarily by dilution rather than other mechanisms) should be included."

Results:

How do the fractional contributions compare between your and Hooper's results? Were the fractional contributions of the fourth and fifth end-members significant compared to the three end-members used by Hooper? How do the end-member distances change with the end-member composition of CHEMMA?

As we stated before, we have limited our analysis to comparing the similarity of the CHEMMA and Hooper end-members.

Figure 4: Is the Shapiro-Wilk test appropriate for this analysis, as it is usually used for normal distribution test?

We agree that the SW test is not an appropriate check on the result, as it only tests whether the overall residual distribution is normal, and not whether there are higher-order relationships between residuals and the independent variables. References to this test and the associated figure have been removed.

Miscellaneous Comments and Suggested Edits

P1/L2-3: This statement is not completely true, not exactly part of the mixing model assumption (see one of my comments above). As long as the temporal/spatial variation is known, a hydrograph separation is still valid.

We have deleted the inaccurate phrase "relatively time invariant".

P1/L17-18: This is not exactly true. The end-member composition does not have to be constant. Delete the phrase.

We have added a word "unknown" to specify the question that we are trying to answer.

P2/L26: Again, it does not have to be temporally invariant.

We have deleted the inaccurate phrases.

P3/L63-65: This is not what Hooper (2003) means. Instead, he demonstrated the filature of using the rule of one. He suggested to use residual distribution pattern. Indeed, their 1992 paper followed the rule of one (as you stated in the sentence following this one). You have to follow the temporal evolution of EMMA and cannot use an early one to reject a later one.

We have changed the order of the statements to better follow the actual timeline.

P3/L69-74: With DTMM, the #2 is not true. The #3 should not be stated as mixing is subject to conservative tracers and so is CHEMMA as I talked above. I do not think CHEMMA can include non-conservative solutes.

We have added a sentence to the end of #2 and changed #2) to "determining the number of significant PC is subjective to some degree, even with the aid of DTMM".

P4/L113: Why not running DTMM before CHEMMA?

DTMM has been applied to this dataset, and the results are discussed in Hooper (2003).

P4/L114: Based on Figure 1, "…projected into the 2D subspace spanned by pair-wise PCs"?

We have added a mathematical expression as an example to illustrate this point to avoid repeating.

P5-6: After reading the text, it is still not clear how both I- and J-vectors were generated. Through an optimization procedure itself?

We have added a section after Line 127 to reflect this comment:

"Because SI may be any possible points within the convex-hull constructed using S, J is needed to force SI to be chosen close to the convex-hull boundary. In practice, I and J are estimated iteratively using an optimization procedure until they converge (Eq. 1 and 2 from Thurau (2011)."

P6/L158-163: Kind of arguments are needed to set the stage for multiple runs. But the exact statements here fit better to discussion.

The paragraph has been updated to better communicate the objective.

P7/L187: You cannot just cite Hooper et al. (1990). DTMM has to run to identify conservative solutes and the number of end-members.

We added a sentence at Line 191 to reflect this comment:

"Later Hooper (2003) suggested the rank of the data (Lower Gauge in Hooper (2003) dataset) is at least 3."

P9/L253-259: Some of them are very much speculative.

Speculative parts of this paragraph have been reduced, so it now reads:

"Failure to conform to these assumptions undermines the validity of the method. For example, rare contributions from an end-member can result in the the dispersion of Cluster 3 in Figure \ref{fig: comparisons}b. Temporal variations of the end-member composition could produce the kind of variations seen in PC 3 in Figure \ref{fig: comparisons}d . Fortunately, CHEMMA itself may be a basis for exploring the effects of time-variability. For example, by partitioning the dataset into time periods (or hydrologic state, etc), the apparent temporal variability of end-members could be explored."

P11/L325-330: These are not conclusion.

This may be a matter of style, but we find it useful to incorporate a brief summary in the conclusions.

Title: Get rid of "observations", which I think is redundant.

Observations has been replaced with "timeseries".

P1/L1: Delete ", and is".

Done

P1/L4: Change "additional measurements" to "samples".

Done

P1/L5 and also L12: I think this article is no longer a technical note.

Done

P1/L16: Delete "profile".

Done

P2/L26: Change “observations’ to “samples”.

Done

P2/L28-30: This is probably where you state all the assumptions of mixing models as I suggested earlier.

Done

P2/L33-35: Any citation(s) for this statement? This is the most important statement to justify your study.

Done

P2/L40: Add “estimate uncertainties of” after “to” if those are true.

Done

P3/L61-63: This is only one of a few criteria used to screen end-members (see Hooper, 2003; Liu et al., 2008 and 2020).

Done

P3/L81-82: Change “allows for identification of” to “aims at identifying”, as by far you have not yet demonstrated if you can.

Done

P3/L84-85: I think what you want to say here is that end-member composition does not have to be distinct for all tracers (assumption iii above).

Done

P4/L95: Change “find” to “determine”.

Done

P4/L115: Add “at each 2D subspace” at the end of the sentence. Then, get rid of the middle sentence if “pair-wise” is added as suggested above.

Done

P5/L (not clear which line but the statement following “Result”): Does x-matrix contain the standardized values or original concentrations? Need to specify.

Done

P5/L (#4 in the table): Change “needed” to “found” because the number is variable and it does not matter for how many to be found.

Done

P5/L (#5 in the table): I am sure both I-vector and J-vector were explained in the text, but in the table their function still needs to be specified so that readers understand what SI-matrix means.

Done

P5/L (#6 in the table): Need to say h and H represent for fractional contributions.

Done

P7-8/L193-215: Most if not all of them should be presented under 3.1.

Thank you for helping us improve the quality of illustration. We answer the changes of the following questions collectively.

Figure 1: What are the red crosses in c? Why are they in different positions from those in b? Also, need to say this uses 3D as an example.

We added a sentence at the end of this caption that “The red crosses are the same extreme points marked in b), but are projected back in the three dimensional PC space.” And we added a part behind the first sentence that “using a three dimensional space with four end-members”.

Figure 2: Do not use the same colors for the PC axis’s of solutes as for the triangles; way too confusing. You used “diamonds” not “squares”. In the case of four end-members, this is not exactly the convex hull but the projected convex hull into 2D PC subspace. For four end-members, it needs three PC subspaces. This should be specified.

We have changed the color for the end-member diamonds to make the visualization clear. We have also corrected the name for the markers and added a sentence that “Note that three dimensional subspace is required for four end-members” at the end of the caption.

Figure 3: Do the colors shown in legend match those points on the plot? I do not see blue and red colors in the legend. Also, the same legend used in all cases regardless of the number of

endmembers. Is it necessary to show all four (maybe five) when you are seeking for only three endmembers?

The reason why those colors seems to not match with the legend is that we used transparent markers, where "the color shade of each cluster reflects the concentration of the vertices at its location". And we added the quoted sentence to the end of the caption.

Figure 4: Cannot be self-explained. Averaged from 100 runs? Residuals of what? What does five-fold cross validation mean? Font size too small; resolution too low; symbol size too small too. Very hard to distinguish Ca and Si curves.

Thank you for pointing out this confusing phrase. We have changed the "Average scalar measures" to "Scalar measures", and increased the font size.

Figure 5: Observed in both stream water and end-members? Do not use "observation" for "stream water" as observed end-members were also shown.

We have changed the original "Observation" x-axis tick to "Stream water". We also modified the caption to make it explain better.

Figure 6: You mean "the same size ..."? Also, change "sample" to "samples". Font size too small.

We have changed the original "Observation" x-axis tick to "Stream water". We also modified the caption to make it explain better.

Figure 7: How were algorithm and data uncertainties calculated? I do not think they were covered in the text.

We have added the original "Observation" x-axis tick to "Stream water". We also modified the caption to make it clearer.

Figure 8: (top) Different number of samples? Or, different sample distributions? (caption) Do they all have the same number of samples?

They have the same number of samples and distribution prior to truncation to the mixing space. We have added a sentence in the caption to demonstrate the same sample size and distribution that "All cases (1 to 6) have the same number of samples (1000 samples) and are normally distributed around the inner center of the grey triangle."

Figure 9: I cannot follow the definition of "percent end-member limited". If samples were generated from three synthetic end-members with constrains of each within the fraction of 0-1 and all summing to 1, how come are some outside of the triangle? Standard deviation or normalized uncertainties? Also need to explain if the normalization is the same as the one shown in an earlier figure. What do components X and Y mean here? PC components? If so, say so!

Percent end-member limited is a way we used to demonstrate how much influence those end-member constraints puts on the random generated samples. We modified some of the sentences to make it clear, and we added an additional sentence to the end of the caption that "'Percent end-member limited' is an intuitive alternation of describing the different degrees of sample dispersion from the inner center of the convex-hull.", to clarify the meaning of the term. We also added "Component X and Y represent the two synthetic measures, X [-] and Y [-], in Figure 8." to explain the meaning of X and Y.

Table 1: Title should be placed above the table (different from figures) unless the journal requires to be placed under. Indicate the number of end-members for each block. What are their fractional contributions? Are the contributions from fourth and fifth end-members significant?

The table has been updated to include the number of end-members. We did not calculate the contributions for the reasons we have listed above at Major Concerns section.

Table 2: Title should be above the table if the journal does not require to be under.

Thank you. We have replaced the table title.