

## **Response to comments from the Referee 2:**

### **Major comments:**

**The main problem with this technical note is the lack of definitions and explanations of variables and notations. Albeit I understand that a technical note is supposed to be short, if the authors want this work to be utilized by other geoscientists and hydrologists, they need to improve the readability of the technical note.**

R: We appreciate the Reviewer's comment in line with another Reviewer's evaluation and have followed his/her suggestions throughout the document. We now included more details on descriptions of equations, variables and notations to improve the readability of the technical note.

**If I understand correctly, this technical note details the methodology used by Correa et al. (2019). It would be great if the authors could be more upfront about this in the technical note.**

R: We are grateful for the comment and we agree, in the new version of the Technical Note it is clearly specified that the methodology developed here is the one applied in Correa et al. (2019b). It now reads: "We illustrate this application on the study case published in Correa et al. (2019b), where the authors presented the uncertainty analysis of sources contributions results [...]".

**The Matlab script is relatively hard to follow and understand, due to insufficient comments and documentation:**

**Naming of Matlab Files: Matlab filenames starting with a number cannot be run by Matlab. Please rename in order to avoid confusion for unexperienced Matlab users.**

**Name of csv file is different in repository than as used in matlab script (5\_data.csv rather than data.csv). Please make consistent.**

**The Matlab Script is not very well documented.**

**Neither the script directly, nor the readme pdf explain the actual inputs (only some cryptic A/B/C/D/M without explanations; one has to refer to Table 1 to understand the setup of the file) nor the outputs.**

**Consequently, the script cannot be used without reading the manuscript in detail.**

**It would furthermore be helpful to reference the code lines to the equations of the manuscript and state in the script what each command is doing.**

**Please also reference the manuscript in the readme.pdf and the main script.**

### **Specific Comments:**

R: We are very grateful that the reviewer highlights the lack of readability in the MatLab script and its documentation. We intended to use the script along with the technical note where the definitions of terms are stated, however, we have included a detailed description in the documentation of the script. Additionally, we have maintained consistency between the repository and the code.

The number of equations in the manuscript have been included in the code (method.m) for reference, along with a description of what each equation does.

The manuscript (under review) will be referenced in the readme file.

Please find at the end of this document, the updated readme and method files.

### **Specific Comments:**

**p.1 line 13: the grammar of this sentence is wrong/the sentence is incomplete.**

R: We have edited the phrase. It now reads: "[...], Bayesian approaches to estimate such source uncertainty only exist only sound methods for two and three sources."

**p.1 lines 15-17: There is no connection between the two sentences, in spite of the "however". Sentence 1 talks about large tracer sets from four water sources (but says nothing about the number of tracers), Sentence 2 says the approach can be generalized to any number of tracers. Please make this clearer.**

R: We have updated this section as suggested: “We illustrate the method to compute individual uncertainties in the calculation of source contributions to a mixture, particularly with an example from hydrology, where a 14-tracer set from water sources and streamflow from a tropical, high-elevation catchment were used. Moreover, this method has the potential to be generalized to any number of tracers in a wide range of disciplines.”

**p.5 line 1-2: the reference of Taylor, 1982 should probably follow directly after “Taylor series approximation”.**

R: We have edited the phrase. It now reads: “[...], the first-order Taylor series approximation (Taylor, 1982) for the variance [...]”

**p.7 line 2: the specification of n=270 is no useful without also specifying the number of different streams sampled.**

R: We thank the reviewer for pointing out a typographical error, the correct number of samples is 257 and the number of streams sampled is 5, the phrase was updated: “Streamwater samples from 5 nested streams were collected weekly from March 2013 to April 2014 (n=257) and at a higher frequency during experimental campaigns.”

**References: Correa et al, 2018, SciTotEnv should actually be Correa et al. 2019.**

R: We updated the reference: Correa, A., Breuer, L., Crespo, P., Célleri, R., Feyen, J., Birkel, C., Silva, C. and Windhorst, D.: Spatially distributed hydro-chemical data with temporally high-resolution is needed to adequately assess the hydrological functioning of headwater catchments, Science of The Total Environment, 651, 1613–1626, doi:10.1016/j.scitotenv.2018.09.189, 2019b.

**Table 1: Footnote: “three- axes”. I believe this should be three-dimensional space.**

R: We have edited the footnote. It now reads: “Coordinates of end-members and stream (mixture) medians in three-dimensional space (U1, U2 and U3). n represents the sample size”

**Table 2: Some of the values provided here do not match up with those calculated by the MATLAB Code. Please verify.**

R: We have verified the result from the MatLab code, and the values presented in table 2 are correct. However, in the Zenodo platform, the data.csv preview shows the data rounded to a decimal and a missing column (Stvd U2), it is necessary to download the.csv file to get the complete data. Updating the code, we will improve the visualization of the data.

## Readme file

These codes estimate the uncertainty of individual end-member (source) contributions to streams (mixture) based on a multi-tracer set in a three-dimensional space.

The method.m code shows step-by-step calculations of partial derivatives, degrees of freedom, t-Student and confidence interval limits for each source fraction.

method.m uses the functions Yx.m and dYzdyx.m for its execution.

A, B, C and D represent the set of sources and M the mixture.

Please refer to Correa et al., (2019) in Correa, A., Ochoa-Tocachi, D. and Birkel, C.: Technical note: Uncertainty in multi-source partitioning using large tracer data sets, Hydrology and Earth System Sciences Discussions, 1–14, doi:<https://doi.org/10.5194/hess-2019-197>, 2019 for a very detailed description of the used notation, equations and variables for this example.

The equation numbers used in the code method.m refer to the corresponding ones in Correa et al., (2019).

### Instructions:

Enter data for the median of end-members and mixture, their standard deviations and sample size, all projected in

the three-dimensional PCA-space (U-space in the referred Technical Note). The file must be named data.csv and follows an order similar to the one presented in this example:

	Median in U1	Median in U2	Median in U3	Stvd in U1	Stvd in U2	Stvd in U3	No. samples	No. samples	No. samples
A	26.25	7.29	7.00	0.46	0.36	0.4	25	25	25
B	0.23	5.47	1.98	0.85	1.29	0.69	33	33	33
C	-2.24	-3.93	3.74	0.55	0.58	0.45	37	37	37
D	-5.38	-6.1	-4.85	0.27	0.56	0.15	36	36	36
M	-0.61	-1.04	0.94	2.07	1.1	0.66	257	257	257

Please do not include rows and column names in the data.csv file, here is used for merely visual purposes of the example.

Run the method.m code.

It is recommended to keep all files in a common directory (method.m; dYzdyx.m; Yx.m and data.csv)

After executing the code, the fractions of the contribution of each end member to the mixture are calculated, as well as the

degrees of freedom and upper and lower limits of uncertainties associated with their contribution.

Note: the order of the inputs (end members) is reflected in the outputs (fractions of contribution and uncertainties).

## method file

```
clear, clc;
format short;
%% Data
%% DATA reads the data.csv file with the input information
DATA = csvread('data.csv');
% Ms
Z = [DATA(:,1:3), ones(5,1)];
Z = transpose(Z);
dataM = Z(:,end);
Z = Z(:, 1:end-1);
% Variance
vZ = transpose (DATA(:,4:6)).^(2);
% Sample QTY
nZ = transpose (DATA(:,7:9));
%% Compute the fractions of sources (A, B, C, D) contribution to the mixture (M) (Eq. 2 and implicitly from Eq. 3)
% x in {A,B,C,D,M} = {1,2,3,4,5}
% y in {delta,lambda,phi} = {1,2,3}
f = zeros(4,1);
DYx = @(y,x) Yx(y,x, Z, dataM);
Num = (DYx(1,5)-DYx(2,5))*(DYx(3,3)-DYx(1,3)) - (DYx(1,3)-DYx(2,3))*(DYx(3,5)-DYx(1,5));
Den = (DYx(1,1)-DYx(2,1))*(DYx(3,3)-DYx(1,3)) - (DYx(1,3)-DYx(2,3))*(DYx(3,1)-DYx(1,1));
f(1) = Num/Den;
f(3) = ((DYx(1,5)-DYx(2,5))-(DYx(1,1)-DYx(2,1))*f(1))/(DYx(1,3)-DYx(2,3));
f(2) = DYx(1,5) - (DYx(1,3)*f(3) + DYx(1,1)*f(1));
f(4) = 1 - (f(1) + f(2) + f(3));
%% Compute the partial derivatives for fA, fC, fB and fD (Eq. 4 and implicitly from Eq. 5 to Eq. 8)
% x in {A, B, C, D, M} = {1,2,3,4,5}
% Y,y in {delta,lambda,phi} = {1,2,3}
%% dfAdyx, presents the partial derivative for fA (Eq. 9)
DdYzdyx = @(Y,z,y,x) dYzdyx(Y,z,y,x,Z,dataM);
dfAdyx = @(y,x) Den^(-2)*( ...
    ((DYx(2,3)-DYx(1,3))*(DdYzdyx(3,5,y,x) - DdYzdyx(1,5,y,x)) + ...
    (DYx(3,5)-DYx(1,5))*(DdYzdyx(2,3,y,x)-DdYzdyx(1,3,y,x)) - ...
    (DYx(3,3)-DYx(1,3))*(DdYzdyx(2,5,y,x)-DdYzdyx(1,5,y,x)) - ...
    (DYx(2,5)-DYx(1,5))*(DdYzdyx(3,3,y,x)-DdYzdyx(1,3,y,x)))*Den - ...
    ((DYx(2,3)-DYx(1,3))*(DdYzdyx(3,1,y,x)-DdYzdyx(1,1,y,x)) + ...
    (DYx(3,1)-DYx(1,1))*(DdYzdyx(2,3,y,x)-DdYzdyx(1,3,y,x)) - ...
    (DYx(3,3)-DYx(1,3))*(DdYzdyx(2,1,y,x)-DdYzdyx(1,1,y,x)) - ...
    (DYx(2,1)-DYx(1,1))*(DdYzdyx(3,3,y,x)-DdYzdyx(1,3,y,x)))*Num);
dfCdyx = @(y,x) ((DYx(1,3)-DYx(2,3))^(-2))*( ...
    ((DdYzdyx(1,5,y,x)-DdYzdyx(2,5,y,x)) - ...
    (DdYzdyx(1,1,y,x)-DdYzdyx(2,1,y,x))*f(1) - ...
    (DYx(1,1)-DYx(2,1))*dfAdyx(y,x))*(DYx(1,3)-DYx(2,3)) - ...
    (DdYzdyx(1,3,y,x)-DdYzdyx(2,3,y,x))*( ...
    (DYx(1,5)-DYx(2,5))-(DYx(1,1)-DYx(2,1))*f(1)));
dfBdyx = @(y,x) DdYzdyx(1,5,y,x) - DdYzdyx(1,3,y,x)*f(3) - ...
    DYx(1,3)*dfCdyx(y,x) - DdYzdyx(1,1,y,x)*f(1) - DYx(1,1)*dfAdyx(y,x);
dfDdyx = @(y,x) -dfCdyx(y,x)-dfBdyx(y,x)-dfAdyx(y,x);
%% Compute the variancefor each end-member fraction, fA, fB, fC and fD respectively (Eq. 10 and Eq. 13)
% x in {A,B,C,D,M} = {1,2,3,4,5}
% y in {delta,lambda,phi} = {1,2,3}
v = zeros(4,1);
for x = 1:5
    for y = 1:3
        v(1) = v(1) + (dfAdyx(y,x)^2)*vZ(y,x);
        v(2) = v(2) + (dfBdyx(y,x)^2)*vZ(y,x);
```

```

v(3) = v(3) + (dfCdyx(y,x)^2)*vZ(y,x);
v(4) = v(4) + (dfDdyx(y,x)^2)*vZ(y,x);
end
end
%% Satterthwaite degrees of freedom for each end-member fraction (Eq. 12 and Eq. 14).
% x in {A,B,C,D,M} = {1,2,3,4,5}
% y in {delta,lambda,phi} = {1,2,3}
g = zeros(4,1);
for x = 1:5
    for y = 1:3
        g(1) = g(1) + (((dfAdyx(y,x)^2)*vZ(y,x))^2)/(nZ(y,x)-1);
        g(2) = g(2) + (((dfBdyx(y,x)^2)*vZ(y,x))^2)/(nZ(y,x)-1);
        g(3) = g(3) + (((dfCdyx(y,x)^2)*vZ(y,x))^2)/(nZ(y,x)-1);
        g(4) = g(4) + (((dfDdyx(y,x)^2)*vZ(y,x))^2)/(nZ(y,x)-1);
    end
end
g = (v.^2)./g;
%% Student's t value(two-tailed) to compute 95% confidence intervals (Walpole et al., 2017)
t = tinv(0.95,g);
%% Compute the upper and lower confidence interval limits for each end-member fraction (Eq. 15)
ulim = min(1,f + t.*sqrt(v));
llim = max(0,f - t.*sqrt(v));
%% Present results:
%% f, fractions of sources (A,B,C,D) contribution to the mixture (M),
%% g, degrees of freedom for each end-member fraction
%% ulim and llim, upper and lower confidence interval limits for each end-member fraction
f
g
ulim
llim

```