



**HAL**  
open science

## Quantification the mixing chemical of sources in a stream from downstream chemical data

José Manuel Tunqui Neira, Gaëlle Tallec, Vazken Andréassian, Jean-Marie Mouchel

► **To cite this version:**

José Manuel Tunqui Neira, Gaëlle Tallec, Vazken Andréassian, Jean-Marie Mouchel. Quantification the mixing chemical of sources in a stream from downstream chemical data. Journée des Doctorants, Mar 2019, Paris, France. hal-04497602

**HAL Id: hal-04497602**

**<https://hal.science/hal-04497602>**

Submitted on 10 Mar 2024

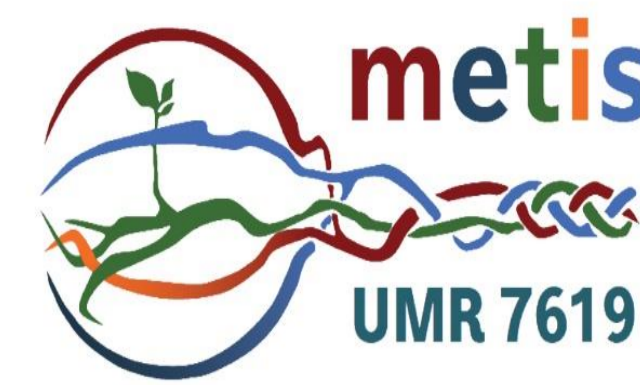
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Quantification the mixing chemical of sources in a stream from downstream chemical data

José Manuel Tunqui Neira<sup>1,2\*</sup>. Gaëlle Tallec<sup>1</sup>. Vazken Andréassian<sup>1</sup>. Jean-Marie Mouchel<sup>2</sup>

<sup>1</sup>UR HYCAR. Irstea. Antony. <sup>2</sup>UMR Metis 7619. Sorbonne University \* [jose.tunqui@irstea.fr](mailto:jose.tunqui@irstea.fr)



## 1. Objective and Context

- ✓ Determining the various sources that mix to produce the composition of a stream from a chemical data set is not an easy task
- ✓ To answer this question and compute the parameters of the mixing model, namely the composition of the sources, our initial hypotheses are:
  1. H1 : the observed concentrations in the stream come from a mixture of  $n$  sources. In our study we consider  $n = 3$ , with 2 components used as tracers that we assume conservative in the stream
  2. H2 : the  $n$  sources are characterized by a stable concentration of each tracer
- ✓ But these hypotheses are not enough to provide an unique solution defining the composition of the sources and a complementary criteria must be added. Any set of sufficiently distant sources, in terms of concentrations, would always produce a perfect decomposition of all data. An additional criteria must force the sources to be not too much distant from each other
- ✓ We have therefore set the following objective function for optimization :

$$F = A * \sum_{i,j} \left( C_{ij} - \sum_k \alpha_{ik} S_{kj} \right)^2 + B * \sum_{k,j} \left( S_{kj} - G_j \right)^2 \quad (1)$$

Where the first term :  $\sum_{i,j} (C_{ij} - \sum_k \alpha_{ik} S_{kj})^2$  is the residual inertia of the concentrations,  $C_{ij}$  (the observed concentration) are a partition of the  $n$  sources  $\alpha_{ik} S_{kj}$  ( $\alpha_{ik}$  the partition et  $S_{kj}$  is the concentration of the source) Additional constraints are:  $\sum_k \alpha_{ik} = 1$ ;  $0 \leq \alpha_{ik} \leq 1$ ;  $S_{kj} > 0$  (2)

Where the second term  $\sum_{k,j} (S_{kj} - G_j)^2$  represents the inertia of the  $n$  sources

- ✓ Since we do not *a priori* information on the best weights (A and B) for the two terms in F, we shall study their effect
- This study aims at defining the right A/B ratio to "best" quantify the sources  $\alpha_{ik} S_{kj}$

## 3. Results

### 3.1 Selection of an « optimal » (A/B) ratio

- Very clear relation between VR and the A/B ratio is observed .The fit curve obtained is a exponential curve
- As an example, we chose an acceptable VR ratio of 0.05 as roughly estimated of the chemical measures errors, which corresponds to a (A/B): **1.23**

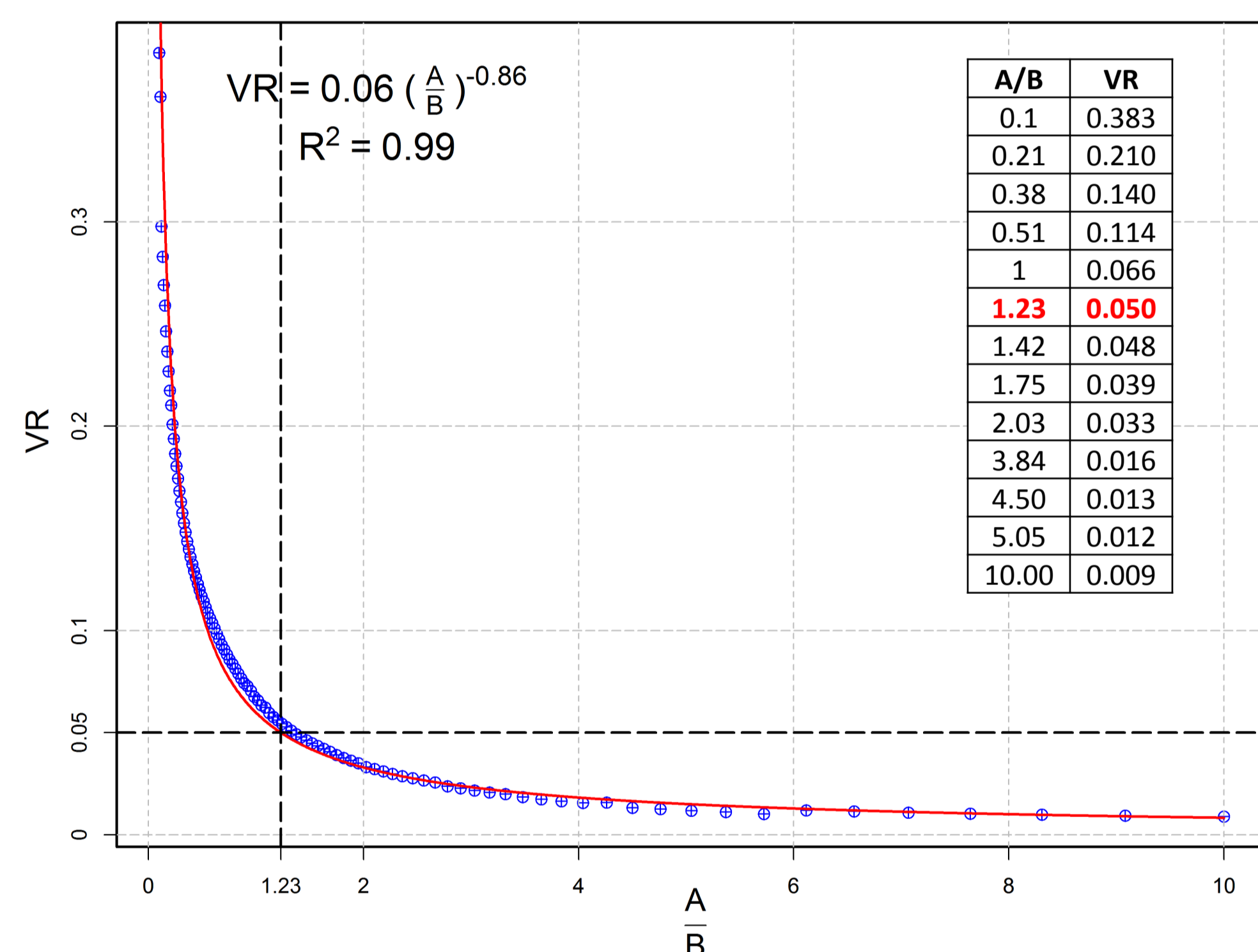


Figure 3 : Fit curve from the scatterplot between (A/B) and VR

### 3.2 Quantification of mixing sources $S_{kj}$

- Our analysis makes it possible to define the sources with respect to the expected VR

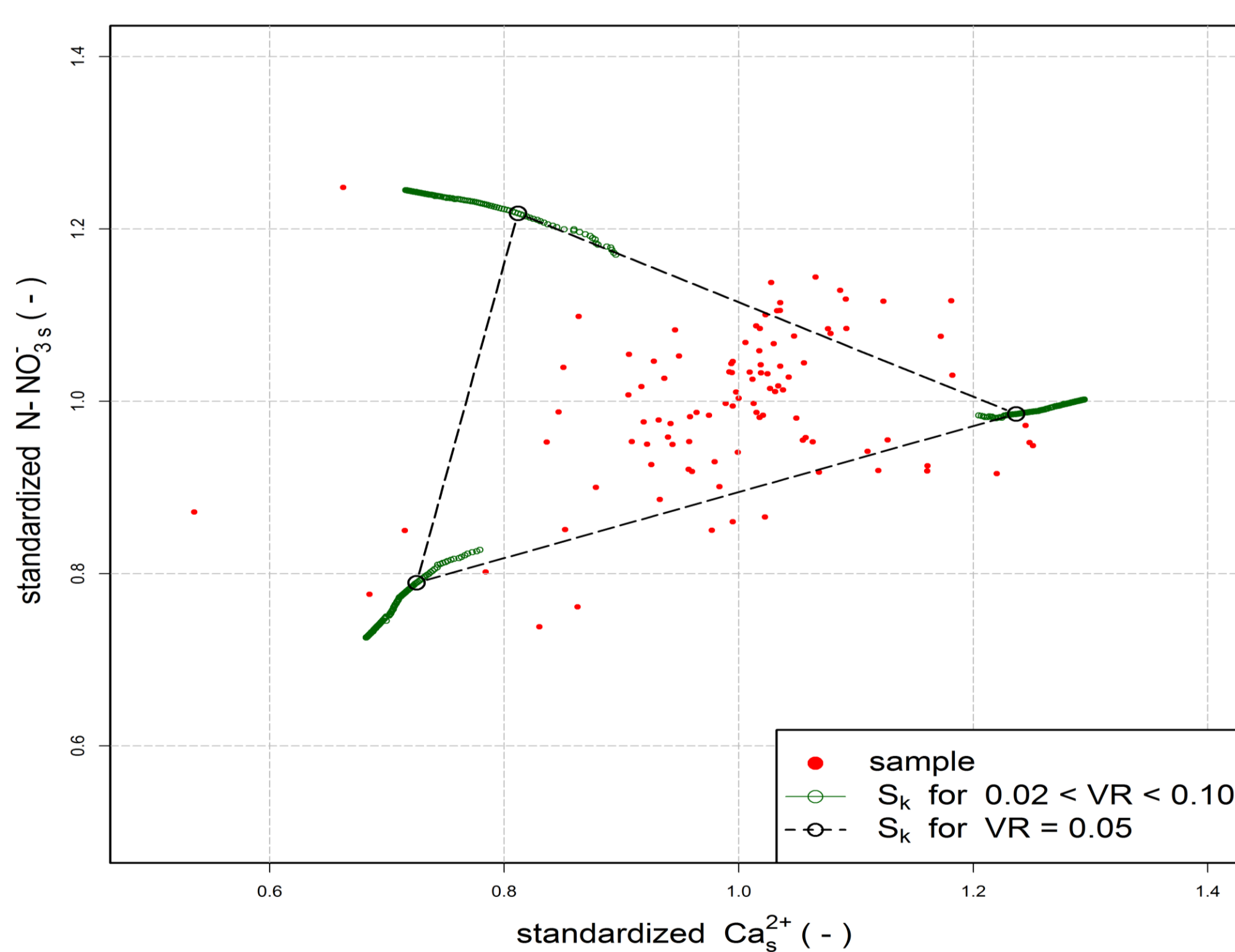


Figure 4 : tracing of the different source compositions  $S_{kj}$  (green lines) calculated for  $0.02 < VR < 0.10$ . The black line represents the source composition for  $VR = 0.05$  and  $(A/B) = 1.23$

- ✓ The 3 different green lines show the evolution of the composition of sources for VR varying between  $0.02 < VR < 0.10$
- ✓ The black triangle show the concentration of the sources obtained for  $VR = 0.05$

Table 2: Source composition obtained for  $(A/B = 1.23$  and  $VR = 0.05)$

Source Composition for $(A/B = 1.23$ and $VR = 0.05)$		
$S_{kj}$	$Ca^{2+}$ (mgL <sup>-1</sup> )	$N-NO_3^-$ (mgL <sup>-1</sup> )
$S_1$	89.18	9.55
$S_2$	99.89	14.74
$S_3$	152.03	11.92

## 2. Material & Methods

### 2.1 Material

- We use high frequency data from Avenelles catchment (ORACLE/Orgeval Observatory) and especially the calcium and nitrate concentrations

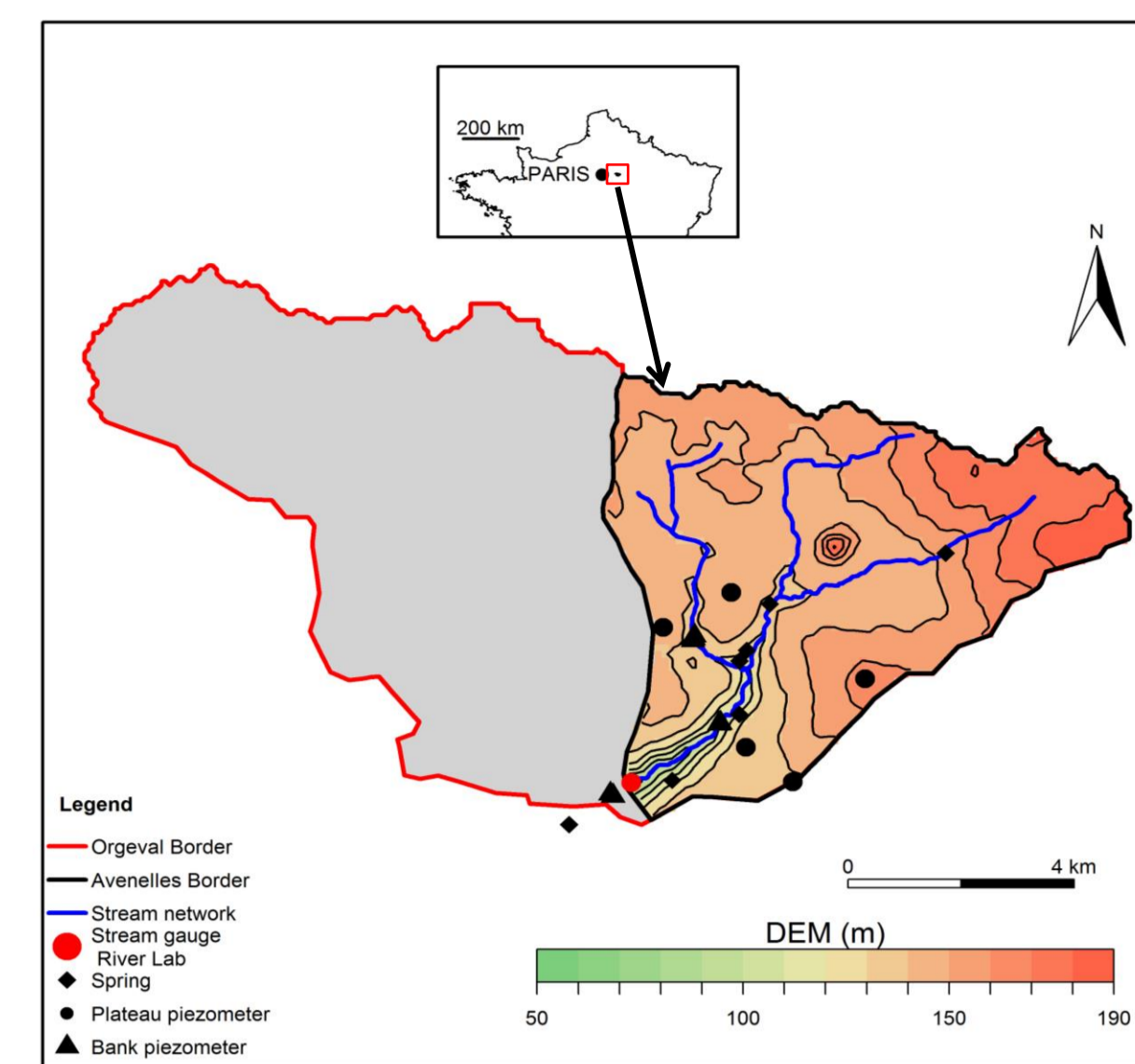


Figure 1 : ORACLE observatory

- ✓ Avenelles catchment (46 km<sup>2</sup>)
- ✓ High-frequency measurements from River Lab [2] (approximately every 30 minutes) from June 2015 to July 2017
- ✓ Seven chemical elements measured (Table 1)

Table 1: Summary of the mean values, min and max of the chemical elements studied from the high frequency measurements

item	Unit	Avenelles Catchment		
		Mean	Min	Max
chloride	mgL <sup>-1</sup>	30	4	40
sulfate	SmgL <sup>-1</sup>	19	2	32
magnesium	mgL <sup>-1</sup>	9	3	15
sodium	mgL <sup>-1</sup>	13	2	17
nitrate	NmgL <sup>-1</sup>	12	3	18
calcium	mgL <sup>-1</sup>	124	37	202
EC	μSm <sup>-1</sup>	704	267	1015

- For computational reasons, we only use a sample of 100 random points from the data series (over 16000 points) (Figure 2, red dots)

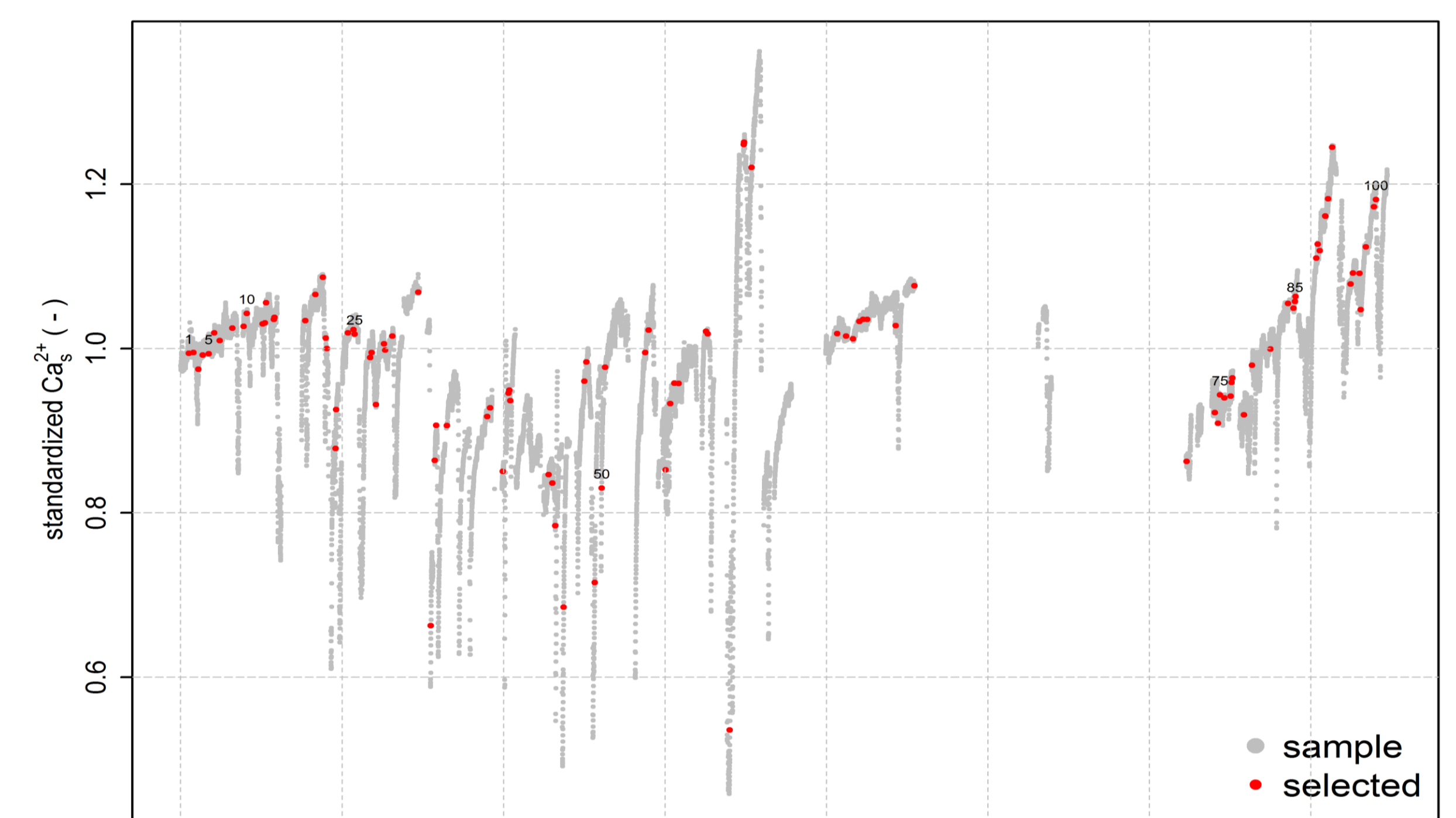


Figure 2 : Points selected (in red) for our sample, taken from our HF data set (gray dots)

- ✓ These 100 points contain most of the different phenomena captured in the high frequency data set
- ✓ Calcium and nitrate concentrations were standardized. i.e. the measured concentrations were divided by their median:

$$c_s = \frac{C_j}{\hat{C}_j}$$

$c_s$  : standardized concentration(-)

$C_j$  : streamflow concentration (mg.L<sup>-1</sup>)

$\hat{C}_j$  : Median of the stream concentration(mg.L<sup>-1</sup>)

### 2.2 Methods

- We applied 100 different A/B ratio ranging from 0 to 10.
- We evaluate the quality of the estimation by the ratio of the residual variances to initial variance of the data set:

$$VR = \frac{\sum_{i,j} (C_{ij} - \sum_k \alpha_{ik} S_{kj})^2}{\sum_{i,j} (C_{ij} - G'_j)^2} \quad (3)$$

Where  $G'_j$  is the center of gravity of the scatterplot of points  $C_{ij}$

- To solve  $\alpha_{ik}$  and  $S_{kj}$  : we used the "solnp" function of the RSOLNP package on R (<https://cran.r-project.org/web/packages/Rsolnp/Rsolnp.pdf>). The "solnp" algorithm is based on the non-linear optimization of the parameters using the augmented Lagrangian method [2]. It makes it relatively easy to deal with both inequality and equality constraints simultaneously (Eq.2).

## 4. Conclusions

- Our model allows to quantify mixing chemical sources in a stream at the cost of an additional constraint regarding the non-explained variance of the concentrations in the stream

## 5. Perspectives

- For the moment, the physical origin of these sources (rain, aquifer, soil, etc.) has not been determined, but the obtained composition of the sources will be compared to what is known about the composition of various potential sources in the catchment.
- Several random extraction in the data set will make it possible to observe the variability of the obtained composition of sources.
- Interesting information on the hydrological behavior of the catchment can be obtained by apply the methodology to the entire record through a sliding window.

## 6. References

- [1] Floury et al. (2017) The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data. Hydrol. Earth Syst. Sci. 21 (2017) 6153-6165.
- [2] Ye, Y., 1988. Interior algorithms for linear, quadratic, and linearly constrained convex programming, Stanford University, 120 pp.

Chemical HF data of the River-Lab are issued from an Irstea-IPGP partnership in the framework of the ANR project CRITEX