Advancing sediment fingerprinting techniques:
The importance of considering sediment mixtures data for tracer selection

Leticia Gaspar [1], Borja Latorre [1], Iván Lizaga [2], Ana Navas [1]

[1] Soil and Water Dept. Aula Dei Experimental Station EEAD-CSIC. Av. Montañana, 1005, 50059, Zaragoza. Spain
[2] Isotope Bioscience Laboratory ISOFYS. Ghent University. Coupure Links 653, 9000, Gent. Belgium

EGU General Assembly 2024
Vienna, Austria & Online | 14–19 April 2024

FingerPro unmixing model

fingerpro@eead.csic.es

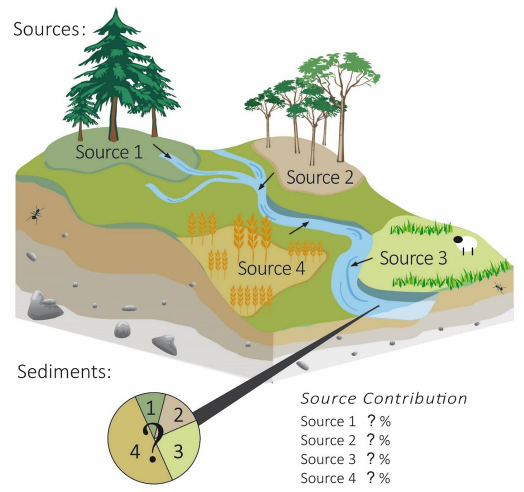
Soil Erosion Group
EESA | EEAD - CSIC
Erosion and Evaluation of Soil and Water

Aula Dei Experimental Station
Soil and Water Dept.
SPAIN

CSIC Consejo Superior de Investigaciones Científicas
EEAD Estación Experimental de Aula Dei

1. INTRODUCTION

Sediment fingerprinting is a key tool to identify the potential sediment sources and quantify the provenance of sediments in catchments



- Tracers are soil properties with unique characteristics for each sediment source that allow discrimination between them
- Statistical tests are used to identify the tracers for each study case but standardised procedures for tracer selection are lacking
- Tracers should be unchanging (conservative) throughout the transport process
- Careful selection of tracers is crucial for obtaining accurate and meaningful results about sediment origin

OBJECTIVE

Emphasize the importance of considering the information provided by each sediment mixture to the selection of tracers before the unmixing

2. METHODOLOGY

We propose novel methods for tracer selection highlighting the importance of:

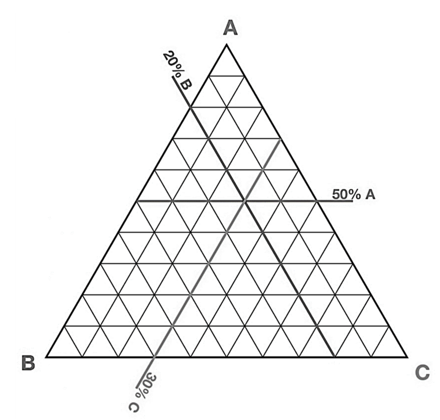
- ✓ Selecting the right tracers for each individual mixture
- ✓ Considering the information on the sediment mixture for the selection of potential tracers
- ✓ Avoid the inclusion of tracers out of consensus or with non-conservative behavior

Range test (RT) are the most widespread methods based on initial mass conservation test, including the information of the sediment mixtures

Kruskal-Wallis and Discriminant function analysis identifies the best combination of tracers that provide the maximum discrimination between sources, but don't include sediment mix information

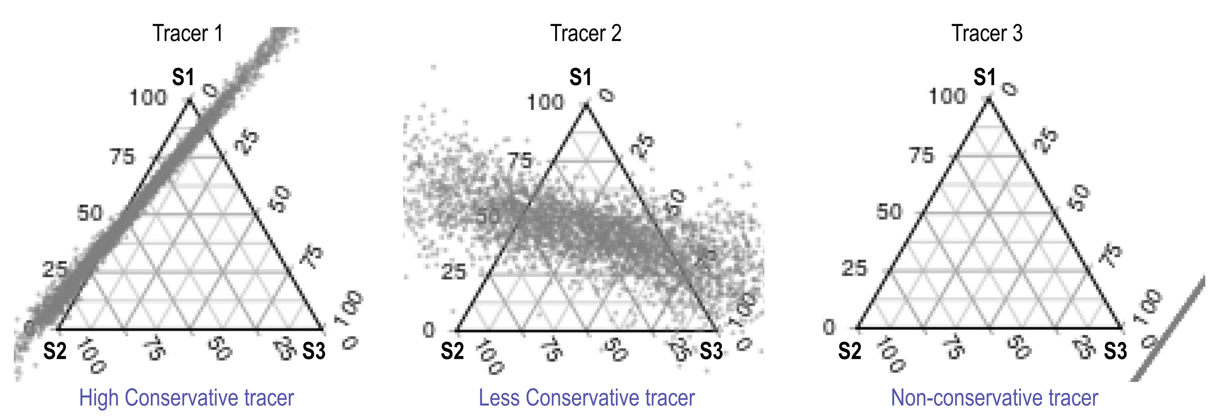
2.1. CONSENSUS METHOD

Conservative Index (CI) analyses individual tracers, identifying the solutions it supports, including the information encoded within the sediment mixture Index that quantify conservative behaviour, *similar to range test*

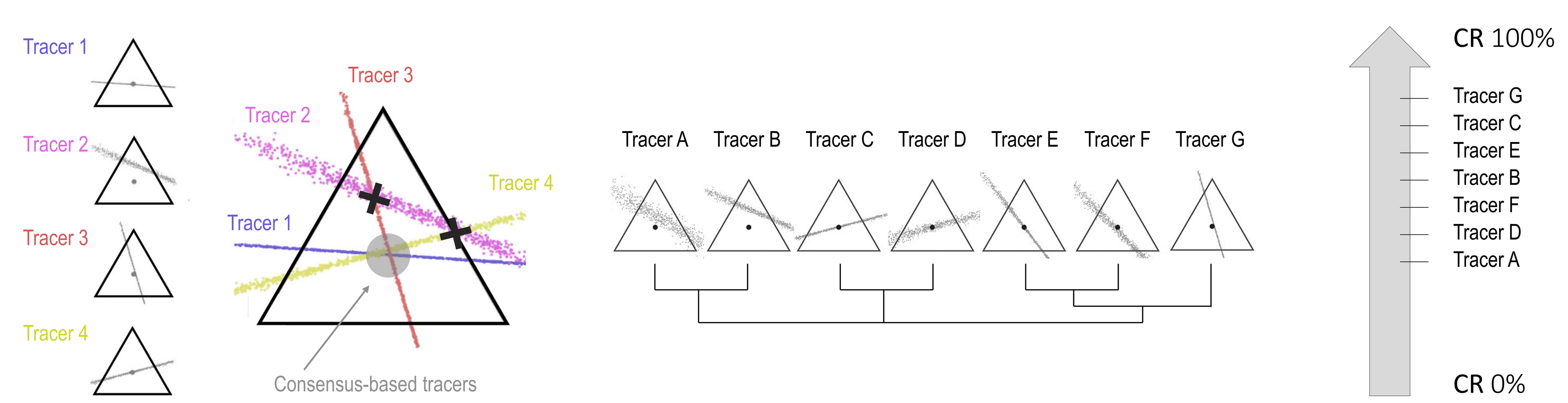


Ternary plots visualise the relative contributions of multiple sources to a sediment mixture. Each corner represent a specific source. Parallel line indicates a different contribution from a specific source. Tracers present line patterns whose direction indicates the source it discriminates

- Points inside the triangle
  All possible combinations
  Source contributions between 0 - 100%
  Mathematical and Physical solutions
- Points outside the triangle
  Impossible combinations
  Source contributions < 0% or > 100%
  Non-physical solutions
- Thickness of the lines
  Indicate the discriminatory power
  Similar to a Wilk's lambda



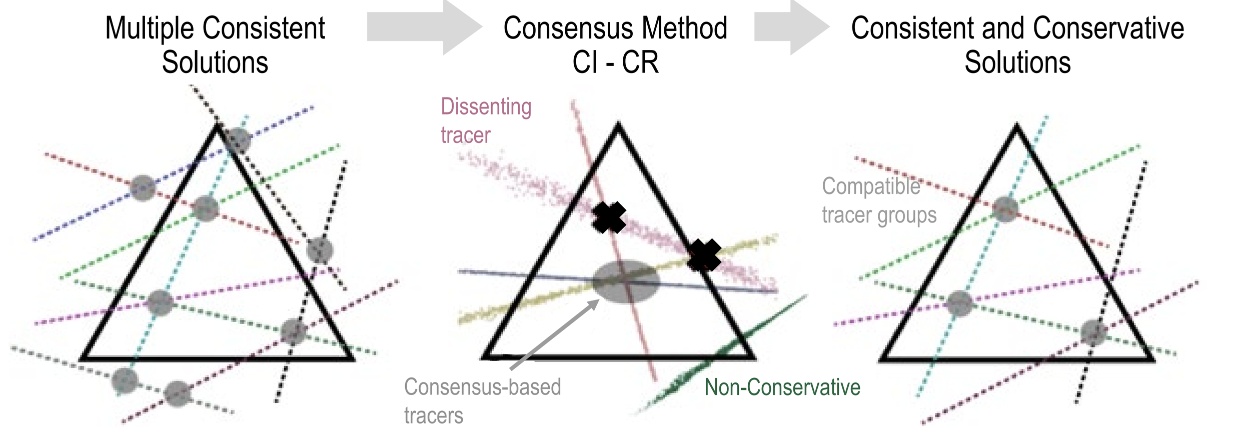
Consensus Ranking (CR) identify conservative tracers (which are highly-ranked) and dissenting tracers. This approach provides mixture-specific insights into tracer behavior



2.2. TRACER COMPATIBILITY

Consistent Tracer Selection (CTS) A novel method for analysing consistency and unravelling multiple solutions in sediment Fingerprinting. CTS method explores the dataset and identifies multiple valid sets of tracers that effectively discriminate sources mathematically consistent (C-test)

The method involved a tracer-by-tracer comparison of the real mixture and its virtual counterpart, with differences standardized by the extreme ranges observed in the sources for each tracer. This method evaluates the residual errors of the model for each tracer

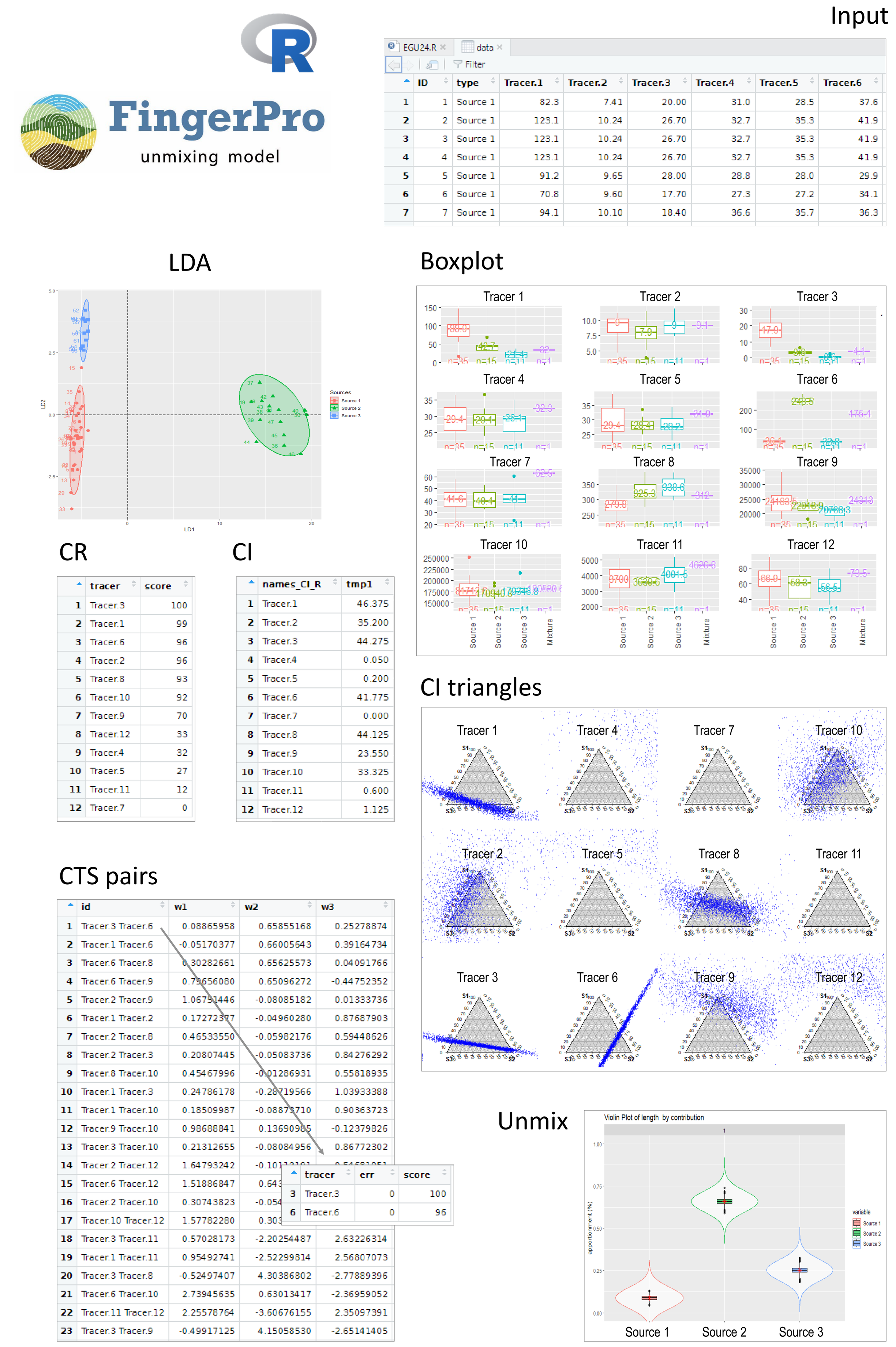


- The lower the residual error for a tracer, the higher the model's confidence in the information carried by that tracer in the mixture.
- CTS builds upon the DFA method. However, DFA focuses on a single, most discriminant tracer selection, CTS explores all possible tracer combinations and incorporates consistency criteria.
- Tracer conservation is also enforced in the resulting selections through the use of previous methods (RT, MP) and consensus

3. RESULTS | FingerPro

Let's explore all tracer selection methods and compare the contribution of sources obtained using the FingerPro model

```r
### 1. SET WORKING DIRECTORY ###
setwd("C:/Users/letic/Downloads/FingerPro Model/R")

### 2. INSTALL FingerPro package from GitHub (.zip file) ###
install.packages('fingerPro.zip', repos = NULL)
library(fingerPro)
library(dplyr)

### 3. INPUT DATA ###
#Load input database a) prepare database before (all columns numbers, no bdl)
data <- read.csv("Input_EGU24.csv", header = T)
View(data)

### 4. EXPLORATORY TESTS ###

# Box plots and a correlation graph of the loaded dataset
boxPlot(data, columns = 1:12, ncol = 3)
correlationPlot(data, columns = 1:12, mixtures = FALSE)

# Linear Discriminant Analysis (tracers discriminate between the  sources)
LDAPlot(data[, c(1:10)], P3D = FALSE, text = FALSE)
LDAPlot(data[, c(1:10)], P3D = FALSE, text = TRUE) #adding point information

# Principal component analysis PCA graph
PCAPlot(data, components = 1:2) # visualize PCA1 vs PCA2

# Kruskal Wallis test (only to explore data)
data <- KWTest(data, pvalue = 0.05) # Do not use to select tracers

# Discriminant Function Analysis (only to explore data)
data <- DFATest(data) # Do not use to select tracers

### 5. TRACER SELECTION METHODS ###

# 5.1 Range test (RT)
data <- rangeTest(data)

# 5.2.1 Conservative Index (CI)
# a) compute the CI index and the individual tracer solutions
CI <- list()
output_CI <- CI_Method(data, points = 4000, Means = F) # Means=T (mean-SD data)
CI <- output_CI [[length(output_CI)]]
write.csv(CI, file = "output_CI.csv")
# b) Ternary plot with the individual tracers solution
Ternary_diagram(output_CI, tracers = c(1:12), n_row = 3, n_col = 4)

# 5.2.2 Consensus Ranking (CR) 3 sources
# a) define sources and mixtures when using RAW data input
sources <- inputSource(data)
mixture <- inputSample(data)
# b) run CR method for 3 sources
output_CR <- cr_3s(source=sources, mixture = mixture, maxiter = 100, seed = 123)
write.csv(output_CR, file = "output_CR.csv")

# 5.3 Consistency-based tracer selection (CTS) 3 sources
# a) compute pairs/triplets (compute pairs for 3 sources)
pairs_CTS <- pairs(source = sources, mixture = mixture, iter = 100, seed = 123)
head(pairs_CTS)
# b) Exploring the pairs for 3 sources
# explore the first pair
sol <- pairs_CTS[1,]
# or explore a  specific pair e.g. Tracer.1 & Tracer.2
sol <- pairs_CTS[pairs_CTS$ID=="Tracer.1 Tracer.2",]
# c) Run CTS
output_CTS <- cts_3s(source = sources, mixture = mixture, sol = c(sol$w1, sol$w2, sol$w3))
# d) select the properties with lower than 5 % of error in CTS and with CR > than 60%
output_CTS <- output_CTS %>% right_join(output_CR, by = c("tracer"))
output_CTS <- output_CTS[output_CTS$err < 0.05 & output_CTS$score > 60,]
head(output_CTS)
write.csv(output_CTS, file = "output_CTS.csv")

### 6. UNMIX sediment mixture ###
# Filter original data to create a dataset with the selected tracers
dataunmix <- dplyr::select(data, "ID","type","Tracer.3","Tracer.6")
View(dataunmix)
# UNMIX with the selected tracers
output_UNMIX <- unmix_R(dataunmix, iter = "short", Means = FALSE)
write.csv(output_UNMIX_R, file = "output_UNMIX_R.csv")
# Plot solutions
# density plot
output_UNMIX_plot <- plotResults(output_UNMIX, y_high = 1, n = 1)
# violin plot
output_UNMIX_plot <- plotResults(output_UNMIX, y_high = 1, n = 1, violin = T)
```



4. TAKE HOME MESSAGE

- ✓ Our findings highlight the importance of considering the information on the sediment mixture for the selection of potential tracers, an aspect often neglected by conventional methods
- ✓ The role of selecting appropriate tracers demonstrates their impact on the results of the unmixing model
- ✓ Include all properties as potential tracers can result in biased findings due to the use of tracers that are either not coherent or not conservative

Authors


Leticia Gaspar
lgaspar@eead.csic.es


Borja Latorre


Iván Lizaga


Ana Navas

References

- Gaspar, L., Blake, W.H., Smith, H.G., Lizaga, I., Navas, A., 2019. Testing the sensitivity of a multivariate mixing model using geochemical fingerprints with artificial mixtures. Geoderma 337, 498–510.
- Gaspar, L., Blake, W.H., Lizaga, I., Latorre, B., Navas, A., 2022. Particle size effect on geochemical composition of experimental soil mixtures relevant for unmixing modelling. Geomorphology, 403, 108178.
- Latorre, B., Lizaga, I., Gaspar, L., Navas, A., 2021. A novel method for analysing consistency and unravelling multiple solutions in sediment fingerprinting. Sci. Total Environ. 789, 147804.
- Lizaga, I., Latorre, B., Gaspar, L., Navas, A., 2020b. FingerPro: an R package for tracking the provenance of sediment. Water Resour. Manag. 34, 3879–3894.
- Lizaga, I., Latorre, B., Gaspar, L., Navas, A., 2020a. Consensus ranking as a method to identify non-conservative and dissenting tracers in fingerprinting studies. Sci. Total Environ. 720, 137537.

Acknowledgment

This research has been supported by projects I+D+i PID2019-103946RJI00 and PID2019-104857RB-I00 funded by MCIN/AEI/ 10.13039/501100011033. L. Gaspar is a Ramón y Cajal researcher at the EEAD-CSIC, funded by MCIN/AEI RYC2020-030338-I