https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Sensitivity of source sediment fingerprinting modelling to tracer selection methods

Thomas Chalaux-Clergue[a], Rémi Bizeul[a], Pedro Batista[b], Núria Martínez-Carreras[c], J. Patrick Laceby[d], and Olivier Evrard[a]

[a]Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL), Université Paris-Saclay, UMR 8212 (CEA-CNRS-UVSQ), Gif-sur-Yvette, France
[b]Water and Soil Resource Research, Institute for Geography, Universität Augsburg, Alter Postweg 118, 86159 Augsburg, Germany
[c]Environmental Research and Innovation Department (ERIN), Catchment and Eco-Hydrology Research Group (CAT), Luxembourg Institute of Science and Technology (LIST), Belvaux, Luxembourg
[d]Alberta Environment and Parks, 3535 Research Rd NW, Calgary, Alberta, T2L 2K8, Canada

Correspondence: Thomas Chalaux-Clergue (thomas.chalaux@lsce.ipsl.fr) - Rémi Bizeul (remi.bizeul@lsce.ipsl.fr)

Abstract. In a context of accelerated soil erosion and sediment supply to water bodies, sediment fingerprinting techniques have received an increasing interest in the last two decades. The selection of tracers is a particularly critical step for the subsequent accurate prediction of sediment source contributions. To select tracers, the most conventional approach is the so-called three-step method, although, more recently, the consensus method has also been proposed as an alternative. The outputs of these two approaches were compared in terms of identification of conservative properties, tracer selection, contribution modelling tendency and performance on a single dataset. As for the tree-step method, several range test criteria were compared, along with the impact of the discriminant function analysis (DFA).

The dataset was composed of tracing properties analysed in soil (through the consideration of three potential sources; n = 56) and sediment core samples (n = 32). Soil and sediment samples were sieved to 63 μm and analysed for organic matter, elemental geochemistry and diffuse visible spectrometry. Virtual mixtures (n = 138) with known source proportions were generated in order to assess model accuracy of each tracer selection method. The Bayesian un-mixing model MixSIAR was used to predict source contributions on virtual mixtures and actual sediments.

The different methods tested in the current research can be distributed into three groups according to their more or less restrictive identification of conservative properties, which were found to be associated with different sediment source contribution tendencies. The less restrictive selections of tracers were associated with a dominant and constant contribution of forests to sediment, whereas the most restrictive selections were associated with dominant and constant contributions of cropland to sediment. In contrast, intermediately restrictive selection of tracers led to more balanced contributions of both cropland and forest to sediment production. Virtual mixtures allowed to compute several evaluation metrics, which supported a better understanding of each tracer selection modelling accuracy. However, strong divergences were observed between the predicted contributions of virtual mixtures and the predicted sediment source contributions. These divergences may likely be attributed to the occurrence of a non-(fully) conservative behaviour of potential tracing properties during erosion, transport and deposition processes, which could not be reproduced when generated the virtual mixtures.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Among the compared tracer selection methods, the three-step method using the mean ± SD and hinge range test criteria provided the most reliable tracer selection methods. In the future, it would be fundamental to generate more reliable metrics to assess conservativeness, to support more reliable modelling and more realistic virtual mixture generation to correctly evaluate modelling accuracy. These improvements may contribute to trustworthy sediment fingerprinting techniques for supporting efficient soil conservation and watershed management.

1 Introduction

During the last several decades, an acceleration of soil erosion has been observed in response to land use changes or farming-practice modifications in several regions around the world (Poesen, 2018; FAO, 2019). Moreover, global warming will likely further increase the frequency of erosive storms and the associated soil losses (OCC, 2015; Li and Fang, 2016). This acceleration of soil erosion leads to an increase of on-site and off-site negative socio-environmental impacts (Lal, 1998, 2001), including the deterioration of soil agronomic properties (Pimentel, 2006; Montgomery, 2007), the transfer of pollutants associated with soil particles (e.g. pesticides, herbicides, chemical fertilizers, heavy metals, radionuclides) (Lal, 1998; Bing et al., 2013; Debnath et al., 2021), the alteration of soil organic carbon stocks (Olson et al., 2016; Lal, 2019), the degradation of aquatic ecosystems (e.g. eutrophication, increased turbidity) (Kemp et al., 2011; Issaka and Ashraf, 2017) and an increased sediment supply to waterbodies (e.g. reservoir and bay siltation) (Collins et al., 2020). The identification of soil erosion sources is therefore essential to prevent water-erosion-induced land degradation and its associated effects.

The sediment source fingerprinting technique was initially developed to determine the origin of sediment (Wall and Wilding, 1976; Peart and Walling, 1986; Loughran et al., 1987). After initial qualitative studies (Wall and Wilding, 1976; Loughran et al., 1987), the subsequent development of quantitative un-mixing models (Peart and Walling, 1986; Walling and Woodward, 1992; Collins et al., 1997a) made it possible to estimate the contributions of different sources to target sediment samples. Since then, the technique has received increasing attention (Collins et al., 2020; Batista et al., 2022). Overall, the goal of sediment tracing studies has been to improve our understanding of sediment transfer processes and to guide landscape management (Laceby et al., 2015; Owens et al., 2016). However, in practice, the technique has mainly been used by scientists as a research tool and few direct applications by landscape managers have been reported (Minella et al., 2008; Collins et al., 2020; Xu et al., 2022). This likely demonstrates that, despite some homogenisation and simplification efforts (Mukundan et al., 2012; Collins et al., 2017; Evrard et al., 2022), the technique remains too complex and the development of simpler and more robust procedures would allow for its wider application (Xu et al., 2022).

In the last few years, there has been a renewed interest among the sediment fingerprinting community in methodological issues associated with the technique, such as the tracer selection methods (i.e., the identification of fingerprint properties suitable for source discrimination and apportionment) (Collins and Walling, 2004; Laceby et al., 2017; Collins et al., 2020;

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Evrard et al., 2022). This might stem from the large diversity of properties that are currently used in sediment fingerprinting studies, for example radionuclides (Collins et al., 1997b; Evrard et al., 2020a), elemental geochemistry (Collins et al., 1997b; Blake et al., 2006; Laceby and Olley, 2015), magnetic susceptibility (Lizaga et al., 2019), organic matter and stable isotopes ($\delta^{13}C$, $\delta^{15}N$) (Laceby et al., 2016b; Huon et al., 2018). Collins et al. (2020) listed properties that have recently gained attention, such as compound-specific stable isotopes (CSSIs) (Gibbs, 2008), environmental DNA (eDNA) (Evrard et al., 2019), the stable oxygen isotope ratio with the oxygen isotopic composition of phosphate ($\delta^{18}Op$) (Mingus et al., 2019), and diffuse reflectance spectroscopy in the visible (Martínez-Carreras et al., 2010; Summers et al., 2011; Tiecher et al., 2015), near-infrared (Summers et al., 2011) or mid-infrared domains (Brosinsky et al., 2014; Farias Amorim et al., 2021). Theoretically, a larger number of measured properties should raise the probability of identifying robust tracers (Laceby et al., 2017; Collins et al., 2020; Evrard et al., 2022). Indeed, tracer selection has a fundamental impact on model predictions and their interpretation (Laceby and Olley, 2015; Laceby et al., 2015; Gaspar et al., 2019), as the inclusion of non-conservative properties in these models was shown to strongly decrease the overall model quality (Sherriff et al., 2015; Smith et al., 2018; Vale et al., 2022).

The most conventional approach of tracer selection is a three-step method (TSM) (Collins et al., 2010; Wilkinson et al., 2013; Laceby et al., 2015; Sherriff et al., 2015). The first step assesses the conservative behaviour, and the second step determines the capacity of discrimination between sources. The results of both tests allow to select tracers from a potentially wide suite of measured properties. The third step of this approach consists in the selection of optimal tracers for modelling.

Conservativeness refers to the absence of changes in the property between sources and targets. Sources correspond to materials that may have contributed to the formation of the target sediments (e.g. soils under different usages, lag sediment, or suspended matter). The nature of the target sediments can vary, as it may include material as different as lag sediment, lake sediments, suspended matter, etc. The non-conservativeness of a tracer can be attributed to two phenomena. The first is that particle size sorting may occur along the transport pathway (Walling et al., 2000). Sediment transport is a physical mechanism which, depending on runoff magnitude and rainfall intensity, will transport specific particle size fractions, weights and natures (i.e. mineral or organic fractions) (Viparelli et al., 2013; Gateuille et al., 2019). In general, the average size of particles decreases with the distance travelled (Laceby et al., 2017). Fine particles with a higher specific surface area are generally associated with higher tracer concentrations than coarser material fractions (Horowitz, 1991; Collins et al., 1997a). In order to reduce the impact of particle size sorting on tracing properties, the $< 63\ \mu m$ fraction is commonly analysed after sieving both source and target material to this threshold (Laceby et al., 2017). The second phenomena is related to tracer concentration changes due to biogeochemical processes occurring during particle transport (Koiter et al., 2013). The changes depend on the tracer tendency to react to biogeochemical processes, such as dissolution, sorption, oxidation and reduction. Highly reactive elements, such as Na, Ca and Mg, show a high water-solubility and tend to dissolve when the sediment is immersed. Other elements, such as Ti, Al, and Si are, in contrast, less susceptible to react in changing conditions, which makes them more suitable tracers (Meybeck and Helmer, 1989; Phillips and Greenway, 1998).

To assess property conservativeness, the conventional tracer selection approach compares the range of property values in source and in target samples. The objective of the range test is to assess whether the range of source values includes all target sediment sample values (Wilkinson et al., 2013). Various range tests based on different statistics are commonly used in the

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

literature: minimum-maximum (Smith and Blake, 2014; Sellier et al., 2021), minimum-maximum ± 10 % (as measurement error) (Gellis and Noe, 2013; Gellis and Gorman Sanisaca, 2018; Dabrin et al., 2021), box-plot examination (whiskers and hinge box) (Sellier, 2020; Batista et al., 2022), mean (Wilkinson et al., 2013; Nosrati et al., 2021), mean plus/minus standard deviation (sd) (Evrard et al., 2020b; Laceby et al., 2021a), and median (Collins et al., 2013; Batista et al., 2022). In these range tests, the source property range is defined as the highest and lowest values of the chosen statistics among the source class. However, range tests do not quantify or confirm the complete absence of any non-conservativeness (Collins et al., 2017; Sherriff et al., 2015).

The property's ability to differentiate between sources, originally proposed by Collins et al. (1997b), determines whether a property is a discriminant tracer or not. This step allows the selection of tracers that maximise the characterisation of a specific source. To assess the discrimination power of a given property, the conventional tracer selection approach relies on the non-parametric Kruskal-Wallis H-test (Hollander, 1973).

Conventionally, after assessing the tracer's conservativeness and discrimination capacity, a discriminant function analysis (DFA) is carried out. Within the DFA, a subset of tracers is selected using a forward stepwise selection procedure based on Wilk's Lambda criterion (Collins et al., 1997b). This step aims at selecting the lowest number of tracers that maximises sample source discrimination, in order to avoid selecting too redundant tracers (Small et al., 2004). However, this practice is currently debated, as some authors argue that a higher number of tracers can reduce prediction uncertainties (Martínez-Carreras et al., 2008; Sherriff et al., 2015).

Another tracer selection method was developed by Lizaga et al. (2020a): the consensus method, which is a procedure based on the information provided by single tracers in an unmixing context. The consensus method selects tracers combining the identification of non-conservative behaviour and conflicting tracers. It consists of two tests: the Conservativeness Index (CI) and the Consensus Ranking (CR). The Conservativeness Index is based on the results of the predictions from single-tracer models to identify non-conservative and dissenting tracers. The Consensus Ranking is a scoring function based on debates aimed at discarding the properties that prevent consensus. Whereas the Conservativeness Index is applied to all target sediment samples and provide unique results for the entire study, the Consensus Ranking is applied to each individual target sediment sample, which may result in the selection of different lists of tracers for different target samples.

Selected tracers are then used in un-mixing models to assess the contribution of sources to the target samples. After the use of simple (Peart and Walling, 1986) and un-optimised quantitative un-mixing models (Collins et al., 1997a), earlier modelling approaches were based on deterministic optimisation procedures (Walden et al., 1997) and, more recently, more advanced approaches have moved towards stochastic procedures using Bayesian or/and Monte Carlo methods (Nosrati et al., 2014; Laceby and Olley, 2015). In order to assess the overall reliability of the study, it is important to assess the predictive accuracy of the un-mixing models. Stochastic models produce a distribution of source contributions for which a prediction interval can be determined, which provides an indicator of modelling accuracy (Batista et al., 2022). The use of artificial mixtures allows prediction accuracy to be assessed in more diverse ways by using them as target mixtures with known contributions (i.e. labelled data). It is then possible to calculate various statistics to describe and evaluate the prediction uncertainty. Although these mixtures were initially prepared in the laboratory (Martínez-Carreras et al., 2010; Haddadchi et al., 2014; Huangfu et al.,

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

2020), the development of virtually generated mixtures (Laceby et al., 2015; Palazón et al., 2015; Sherriff et al., 2015) appears as a relevant alternative (Batista et al., 2022). However, when artificial mixtures are produced, their properties are not affected by erosive processes and are therefore perfectly conservative. The selection of tracers is therefore crucial when assessing the accuracy of modelling using artificial mixtures to truly reflect reality.

The objectives of the current study are therefore to: (1.a) compare the tracers selection given by two tracer selection approaches (i.e. three-step method and consensus method), (1.b) assess the impact of the stepwise selection on the three-step method, (2) evaluate the impact of these different tracer selections on sediment source apportionment prediction accuracy using virtual mixtures and (3) draw general recommendations from this evaluation for future sediment fingerprinting studies.

2 Materials and Methods

2.1 Catchment description

The Hayama lake catchment ($84\ km^2$), located in the upper part of the Mano River in Northeastern Japan (Fukushima Prefecture, Tohoku Region), is a typical mountainous agricultural catchment of the eastern edge of the Fukushima prefecture. Due to the steep topography, cropland is located at the bottom of valleys and in the vicinity of rivers, and it is bordered by forest on steep mountainous hillslopes. Forestry is the main land use, which covers 91% of the catchment, while cropland and urban settlements represent respectively 7% and less than 1% (Fig. 1 data from JAXA (2016, 2018, 2021)). However, cropland is located in places with a high hydro-sedimentary connectivity (Chartin et al., 2013).

Catchment altitude ranges from $170\ m$ to $700\ m$ above sea level. The climate is continental (Dfa), with no dry season and hot summer, and bordered to the east by a Cfa temperate climate with no dry season and hot summer according Köppen's climatic classification (Beck et al., 2018). The regional hydrological year runs from November to October (Laceby et al., 2016a; Whitaker et al., 2022). Over the 2006 to 2021 period, the mean annual temperature was $13.6 \pm 0.4\ °C\ y^{-1}$ (standard deviation), with mean monthly values ranging from $-1.5\ °C$ in January to $31.1\ °C$ in August. The mean annual precipitation was $1220 \pm 189\ mm\ y^{-1}$ (sd). The majority of precipitation occurs between June and October, representing 60% of the annual rainfall and 86% of the rainfall erosivity (Laceby et al., 2016a). This period corresponds to the Japanese typhoon season with a peak of intensity in September in the Fukushima Prefecture. Major typhoons were shown to be the main drivers of sediment production, (Chartin et al., 2017) as they can generate 40% of the annual rainfall erosivity within a very short period (Laceby et al., 2016a).

The Hayama lake catchment is mainly underlied by non-alkaline mafic volcanic rocks (42%), granite (31%) granodiorite (17%) and sedimentary rocks (7%) (supplementary materials S1). Main soil groups, according to the Comprehensive Soil Classification System of Japan (Obara et al., 2011, 2015, and equivalent soil types according to the World Reference Base for Soil Resources (WRB) were provided) are: Brown Forest soils (37%; WRB: Cambisols/Stagnosols), Allophanic Andosols (36%; Silandic Andosols), Cambic Red-Yellow soils (9%; WRB: Cambisols) and Lithosols (9%; WRB: Leptosols) (Supplementary Materials S2) (data from NARO (2011)).

All maps and geographical processing were performed using the QGIS software (QGIS Development Team, 2022).

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

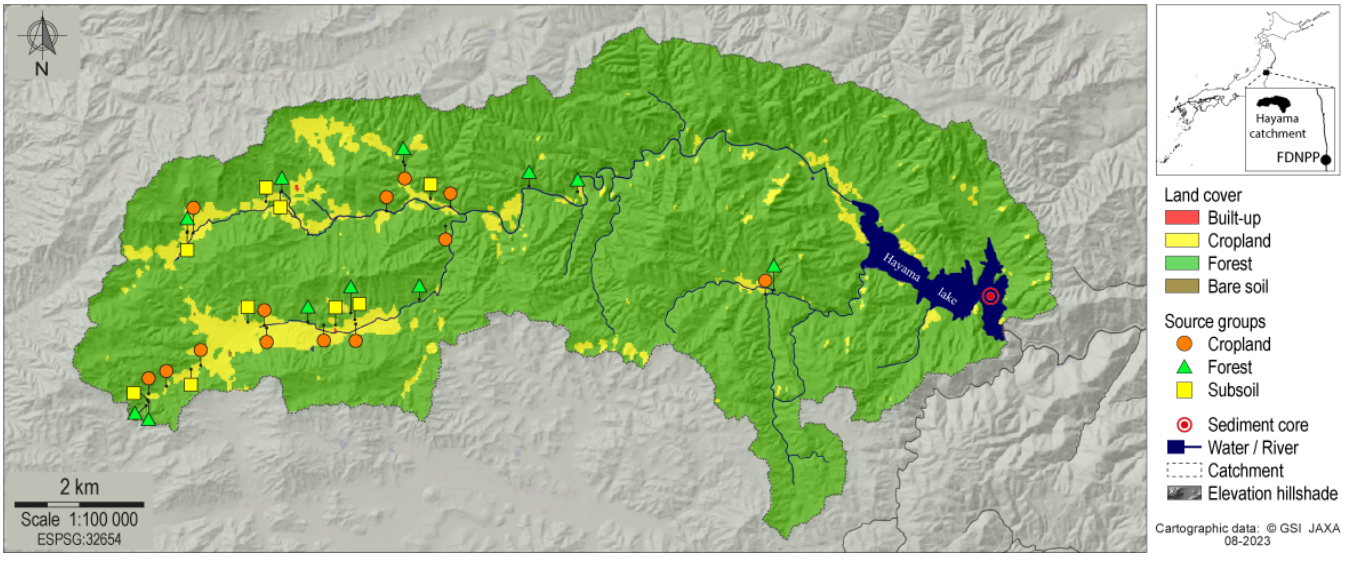

Figure 1. Map of the main land uses in the study area over the 2014-2016 period with location of source samples and the sediment core (cartographic data: GSI and JAXA). FDNPP: Fukushima Dai-ichi Nuclear power plant.

2.2 Soil and sediment sampling and processing

All sediment targets were taken from one sediment core sampled on June 8th, 2021 in the downstream part of Hayama lake (Mano dam lake) at 42 m depth (Fig. 1) by the National Institute of Environmental Studies (NIES) (Japan). The core was 38 cm long with a diameter of 11 cm. It was sectioned into 38 increments of 1 cm. A group of 32 layers was selected for this study, from 6 cm to 38 cm depths, which corresponds to a stable land use period. Sediment samples were dried at 40 $°C$ for 96 hours.

Soil samples (n = 56) were collected in areas representative of the main potential sediment sources to Hayama lake, including 24 cropland, 22 forest and 10 subsoil (i.e. channel bank or land slide) samples. Some source soils were sampled in the adjacent Niida River catchment, which is similar to the Hayama lake catchment. A particular care was taken to ensure that these samples were representative of the Hayama lake catchment characteristics, in terms of land use, geology and pedology (see in Supplementary Materials S1, S2 and S3). Soils were sampled with a plastic trowel and consisted of 10 composited sub-samples of topsoil (1-2-cm uppermost layer). All soil samples were dried at 40 $°C$ for about 48 hours and then successively sieved to 2000 μm and 63 μm.

2.3 Laboratory analysis

Various analyses were conducted to characterise sediment and soil properties: organic matter composition determined by the combustion method (i.e. total organic carbon (TOC) and total nitrogen (TN)), elemental geochemistry analysed by X-ray fluorescence (XRF) for 17 elements (i.e. Al, Ca, Co, Cr, Cu, Fe, K, Mg, Mn, Ni, Pb, Rb, Si, Sr, Ti, Zn, and Zr), and visible

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

colour indices by diffuse reflectance (i.e. CIE Lab, CIE LCh, geothite peaks intensity (445 and 525 nm (Tiecher et al., 2021))), the ratio between the % reflectance at 700 nm and 400 nm (Q7/4 (Debret et al., 2011)) and iron oxide-associated parameters (A1, A2, A3, Gt (Tiecher et al., 2015)). All analysis methods and calculations are described in the supplementary materials.

Only properties for which the measurement uncertainty is too large were removed from further analysis. The following criterion has been set: if when considering the measurement uncertainty a majority or the totality of the samples values were virtually impossible, then the corresponding property was discarded.

2.4 Virtual mixtures

Virtual mixtures were generated in order to assess model prediction accuracy (Palazón et al., 2015; Smith et al., 2018; Gaspar et al., 2019; Farias Amorim et al., 2021; Batista et al., 2022). They were shown to provide a reliable alternative to laboratory mixtures (Batista et al., 2022). Their ease of generation allows the model to be evaluated under a wide range of source contributions. For a given source, virtual mixture contributions were designed to range from 0 to 100%, with 5-percent increments. The contributions of the other sources to the mixtures were then determined as fractions of the remaining contribution (i.e. $1 - source\ A\ contribution$), the fractions being in turn determined by the number of sources. The denominators were defined as: $(number\ of\ sources - 1) * 2$, and the numerators were set to 3 and 1 as three groups were considered in the current research (i.e. $3/4$ and $1/4$). Permutations were then determined following this contribution scale, which generated a total of 138 virtual mixtures for the three source groups. For every source group, the mean and standard deviation (sd) of each property were calculated. For each virtual mixture, each source group mean value was multiplied by the corresponding group contribution (e.g. cropland 0.40, forest 0.45, subsoil 0.15). Then, all source group values were summed to provide the virtual mixture property value.

2.5 Tracer selection

2.5.1 Three-Step Method

The three-step method (Mukundan et al., 2010; Sellier, 2020; Batista et al., 2022) is based on three steps: (1) a range test to identify conservative properties (Martínez-Carreras et al., 2010; Wilkinson et al., 2013; Gellis and Walling, 2013), (2) a Kruskal-Wallis H-test to identify discriminant properties (Collins et al., 1997b). Then, (3) a discriminant function analysis (DFA) forward stepwise selection based on Wilk's Lambda criterion to identify the best subset of predictors among the identified tracers (i.e. conservative and discriminant properties) is run (Collins and Walling, 2002).

Conservativity.

The conservative behaviour of properties is assessed using a range test. Within a range test, the range of sources is defined as the highest and lowest values of a criterion of the property among source groups. To be conservative, all the sample property values should lie within the source range.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

conservative if $\{ \; sources\ lower\ bound \leq target\ value \leq sources\ upper\ bound \}$ (1)

Several range test criteria are found in the literature, these were all applied to identify conservative properties as an alternative to the three-step method. The most common range test criteria were: minimum-maximum (Smith and Blake, 2014; Sellier et al., 2021), minimum-maximum ± 10 % (as measurement error) (Gellis and Noe, 2013; Gellis and Gorman Sanisaca, 2018; Dabrin et al., 2021), boxplot whiskers interpretation (i.e. outlier thresholds) (Sellier, 2020) and boxplot hinge interpretation (i.e. also referred to as interquartile range (IQR)) (Batista et al., 2022), mean (Wilkinson et al., 2013; Nosrati et al., 2021), mean plus/minus one standard deviation (Mean ± SD) (Evrard et al., 2020b; Laceby et al., 2021a), and median (Collins et al., 2013; Batista et al., 2022) are detailed in Table 1.

Table 1. Equations of the common range test criteria used in the literature to test the conservativity in the three-step method. With s_i the source group i from 1 to n the number of source groups, t_j the target sample j value with j from 1 to m the number of samples, and μ the mean and σ the standard deviation (SD). * indicates statistics calculated on log transformed values.

Criterion	Range test equation
Minimum-Maximum	$min(min_{s_i}) \leq t_j \leq max(max_{s_i})$
Minimum-Maximum ± 10%	$min(min_{s_i} * 0.9) \leq t_j \leq max(max_{s_i} * 1.1)$
Hinge	$min(Q_{s_i}(0.25)) \leq t_j \leq max(Q_{s_i}(0.75))$
Whiskers	$min(max(min_{s_i}, Q_{s_i}(0.25) - 1.5\ \text{IQR}_{s_i})) \leq t_j \leq max(min(max_{s_i}, Q_{s_i}(0.75) + 1.5\ \text{IQR}_{s_i}))$
Mean	$min(\mu^*_{s_i}) \leq t_j \leq max(\mu^*_{s_i})$
Mean ± SD	$min(\mu^*_{s_i} - \sigma^*_{s_i}) \leq t_j \leq max(\mu^* s_i + \sigma^*_{s_i})$
Median	$min(median_{s_i}) \leq t_j \leq max(median_{s_i})$

Discriminant power.

The ability of the property to discriminate between source groups is assessed using a Kruskal-Wallis H-test (Hollander, 1973). This is a non-parametric test that checks whether two or more samples originate from the same distribution, and it represents an extension of the Mann-Whitney U-test that compares two samples. Test hypotheses are:

- The null hypothesis (H_0): the median across the three groups are equal.
- The alternative hypothesis (H_1): At least one of the group median is different from the others.

The test statistic is given by:

$$H = (N-1)\frac{\sum_{i=1}^{g} n_i \, (\overline{r}_i - \overline{r})^2}{\sum_{i=1}^{g} \sum_{j=1}^{n_i} (r_{ij} - \overline{r})^2} \quad (2)$$

with N the total number of observations across all groups, g the number of groups, n_i the number of observations in group i, r_{ij} the rank of observation j from group i, $\overline{r}_i$ the mean rank of all observations for group i and $\overline{r}$ the mean of all r_{ij}.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Kruskal-Wallis H-tests were performed using the function kruskal.test from the stats package (R Core Team, 2022). For properties with a p-value below α = 0.5, the null hypothesis is rejected, which means that at least one of the source groups is different from the others and that the property is therefore discriminant.

Discriminant Function Analysis.

A discriminant function analysis (DFA) stepwise selection is carried out to identify a subset of predictors among the identified tracers (Collins and Walling, 2002; Collins et al., 2010). A forward stepwise variable selection based on Wilk's Lambda criterion was performed using the function greedy.wilks from the klaR package (Weihs et al., 2005). The function first initiates a model with the variable that discriminates groups. Then, the model is extended by including other variables based on Wilk's Lambda criterion: the one variable that minimises the Wilk's Lambda is only included if the model's p-value remains statistically significant after its inclusion. Wilk's Lambda statistic tends to zero when the variability between groups is higher than the variability within each group, then this criterion maximizes the difference between groups. As the use of the DFA stepwise selection has been criticised because some studies showed that the use of a higher number of tracers decreases the sensitivity of the results to non-conservative tracers (Martínez-Carreras et al., 2008; Sherriff et al., 2015), we used the list of identified tracers before ("no DFA") and after the stepwise selection procedure ("DFA") in order to assess its potential impact on the calculation of source contributions.

2.5.2 Consensus Method

Lizaga et al. (2020a)'s approach to select properties, referred to as the Consensus Method (CM), is based on two tests: the Conservativeness Index (CI) and the Consensus Ranking (CR). In the CM, tracers are selected by identifying non-conservative behaviour (using the CI test) and dissenting tracers (CR test). The CI of each property is calculated for all target samples simultaneously, while the CR is calculated for each target sample independently. As a result, a set of tracers is selected for each target sample. Tests were performed using the 1.3 version of FingerPro (Lizaga et al., 2020b) under R ver. 4.1.2 (R Core Team, 2021) environment.

Conservativeness-Index.

The Conservativeness-Index (CI) quantifies how conservative a property is based on the result of a single-tracer model. A single-tracer model is a standard linear un-mixing model with only one true analysed property and n-2 (n the number of source groups) virtual tracers. The model is solved to obtain the source contributions. This process is repeated 2000 times, which creates a distribution of predictions. The prediction couples (i.e. $w_1, w_2 ... w_n$) are sorted according to the Euclidean distance to a perfectly balanced mixture (i.e. $\frac{1}{n}$). A percentile of the sorted prediction couple is chosen to compute the CI as the root mean square error (RMSE) of the non-conservative part (nc) of the contribution, as follows:

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

$$\mathrm{CI} = -\sqrt{\sum_{j=1}^{n}\left(\mathrm{nc}(w_j) - \frac{1}{n}\right)^2} \text{ ; with } \mathrm{nc}(w) = \begin{cases} -w & \text{if } w < 0 \\ 0 & \text{if } 0 \leqslant w \leqslant 1 \\ w-1 & \text{if } w > 1 \end{cases} \quad (3)$$

with w_j of the jth source group predicted contribution of the selected percentile, and n the number of source groups. A property is considered as conservative if the CI values is strictly equal to zero.

Consensus Ranking.

Consensus ranking (CR) is an index that quantifies the relevance of each property for the prediction based on a debate agreement for a single target sample. In each debate, a subset of $n+1$ randomly selected properties is built and several rounds of debates are performed excluding one property each time. The consensus of each debate is measured through the quality of the mass balance equations. When the exclusion of a property leads to an increase of the subset consensus score (i.e. higher RMSE of the mass balance equations), the property is considered as dissenting and "lost a debate". The CR of a property is the ratio between the number of attended debates (set at 2000) and the number of lost debates, and it is calculated as follows:

$$\mathrm{CR} = 100\left(1 - \frac{\text{lost debates}}{\text{attended debates}}\right) \quad (4)$$

Properties with a CR score above a certain threshold are considered relevant and are selected. Lizaga et al. (2020a) recommended to select properties with a CR score above 70.

2.6 Source contribution modelling

2.6.1 Un-mixing model

An un-mixing model was run on actual sediment samples and virtual mixtures using the tracers selected by the three-step method (TSM) without and with DFA, and the consensus method (CM). To do so, the widely used Bayesian un-mixing model, MixSIAR, was employed using the R package MixSIAR (Stock et al., 2020, ver. 3.1.12) with JAGS (Stock et al., 2022, ver. 4.3.1) (Collins et al., 2020; Evrard et al., 2022; Lizaga et al., 2020a; Batista et al., 2022). As source and mixture (i.e. sediment samples or virtual mixtures) tracer data are assumed to be normally-distributed in MiXSIAR, tracers were log-transformed to enforce a higher degree of normality prior to modelling (Laceby et al., 2021b). The model was run with a "long" Markov Chain Monte Carlo sampling algorithm (i.e. Chain length = 300,000, Burn-in = 200,000, Thin = 100 and Chains = 3) with a process error structure. Model convergence was determined by the Gelman-Rubin diagnostic using the output_JAGS function from the MixSIAR package, and none of the tracer selection approaches tested had a value greater than 1.05. The median value of the distribution predicted by MixSIAR was reported as the source contribution for each sediment target and virtual mixture.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

2.6.2 Accuracy assessment

For each tracer selection approach, MixSIAR prediction accuracy was assessed using virtual mixtures contribution predictions, for which theoretical source contributions were known a priori. Model's prediction accuracy was evaluated based on different criteria: uncertainty (prediction interval width (W50)), residual error or bias (Mean error (ME)), performance (Squared Pearson correlation coefficient (r^2), root-mean-square error (RMSE), Nash-Sutcliffe modelling efficiency coefficient (NSE) and continuous ranked probability score (CRPS)). A summary of the metrics, with their formula, unit, range and ideal values is provided in Table 2 (Matheson and Winkler, 1976; Bennett et al., 2013; Batista et al., 2022).

Table 2. Formula of model prediction accuracy metrics. With z_i and $\hat{z}_i$, respectively, the measured and predicted contributions for the sample i, $\bar{z}$ and $\bar{\hat{z}}$ the measured and predicted mean contribution of all samples, and n the number of samples.

Name	Unit	Formula	Range	Ideal value
Prediction Interval width (W50)	%	$\mathrm{W50} = Q(0.75) - Q(0.25)$	(0, 100)	0
Mean Error (ME)	%	$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(z_i - \hat{z}_i)$	$(-\infty, +\infty)$	0
Root Mean Square Error (RMSE)	%	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}$	$(0, +\infty)$	0
Squared Pearson's correlation coefficient (r^2)		$r^2 = \left(\frac{\sum_{i=1}^{n}(z_i-\bar{z})-(\hat{z}_i-\bar{\hat{z}})}{\sqrt{\sum_{i=1}^{n}(z_i-\bar{z})^2}\sqrt{\sum_{i=1}^{n}(\hat{z}_i-\bar{\hat{z}})^2}}\right)^2$	(0, 1)	1
Nash-Sutcliffe modelling efficiency coefficient (NSE)		$\mathrm{NSE} = 1 - \frac{\sum(z_i-\hat{z}_i)^2}{\sum(z_i-\bar{z_i})^2}$	$(-\infty, 1)$	1
Continuous ranked probability score (CRPS)		$(F_i, z_i) = \int_{-\infty}^{+\infty}(F_i(z_i) - H\{z_i \leq \hat{z}_i\})^2$	$(0, +\infty)$	0

Higher values of W50 indicate a wider distribution, which is related to a higher uncertainty. The sign of the ME indicates the direction of the bias, i.e. an over- or underestimation (positive or negative value, respectively). As ME is affected by cancellation, a ME of zero can also reflect a balanced distribution of predictions around the 1:1 line. Although this is not a bias, it does not mean that the model outputs are devoid of errors. The RMSE is a measure of the accuracy and allows to calculate prediction errors of different models for a particular dataset. RMSE is always positive and its ideal value is zero, which indicates a perfect fit to the data. As RMSE depends on the squared error, it is sensitive to outliers. The r^2 describes how linear the prediction is. The NSE indicates the magnitude of variance explained by the model, i.e. how well the predictions match with the observations. A negative RMSE indicates than then mean of the measured values provides a better predictor than the model. The joint use of r^2 and NSE allows a better appreciation of the distribution shape of predictions and thus facilitates the understanding of the nature of model prediction errors. The CRPS evaluates both the accuracy and sharpness (i.e. precision) of a distribution of predicted continuous values from a probabilistic model for each sample (Matheson and Winkler, 1976). The CRPS is minimised when the observed value corresponds to a high probability value in the distribution of model outputs. The formulae and the full description of this score are available in Jordan et al. (2017) and Laio and Tamea (2007). The calculation of the CRPS was performed using the crps_sample function from the scoringRules package (Jordan

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

et al., 2019). Model global CRPS value is calculated as the mean of individual CRPS values. In addition to the metric values, a graphical evaluation of model predictions was performed through plotting observed versus predicted, CRPS sample values and W50 sample values for each source group.

All data analyses were performed using R (R Core Team, 2022, ver. 4.2.2) within RStudio (RStudio Team, 2022). An R package (fingR) implementing the approach followed in this study was developed and is freely available (Chalaux-Clergue and Bizeul, 2023).

3 Results

According to the measurement uncertainty criterion (s.2.3), the following properties were removed from subsequent analysis: the elemental concentrations in Co, Cr, Cu, Ni and Rb; the visible colorimetric index A3, and the goethite peak at $445\ nm$ (G_{445}), as their measurement uncertainties were too high.

3.1 Tracer selection

3.1.1 Selection of tracers

Three-step method.

The different range tests resulted in unique sets of conservative properties, with only TN and Q7/4 passing all tests (Fig. 2). Nevertheless, the majority of tests identified TOC, G_{525} b*, Al, Ti, L* and C*. The minimum-maximum $\pm$ 10 $\%$ is the range test criterion that identified the highest number of properties (19), followed by the minimum-maximum (16) and whiskers (12). The minimum-maximum $\pm$ 10 $\%$ range test criterion was the only one that identified Fe, Pb, h, A1 and GT. As the minimum-maximum criterion, this test identified Ca, K, Mg and Sr as conservative. The mean and median criteria identified only three (TN, b*, Q7/4) and four properties (TOC, TN, Q7/4, G_{525}), respectively. Among the properties identified as conservative, Fe, Mg, Pb and Ti were identified as non-discriminant by the Kruskal-Wallis H-test (p-value = 0.08) and were therefore removed from the list of potential tracers. Although a majority of tests only identified Ti as conservative, only the minimum-maximum $\pm$ 10 $\%$ criterion identified Fe, Mg and Pb. The stepwise selection procedure mostly resulted in the systematic exclusion of A2, and in the frequent exclusion of TN and G_{525} while TOC, Al, Si, L*, and Q7/4 were retained by most of the range test criteria. Among the 16 tracers selected by the minimum-maximum criterion, the stepwise selection discarded seven of them (TN, Ca, Sr, b*, C*, A2 and G_{525}). In contrast, the minimum-maximum $\pm$ 10 $\%$ criterion selected 19 tracers and discarded only five tracers: TOC, TN, Ca, A2 and G_{525}. The outputs of the mean and median criteria were not modified by the stepwise selection.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

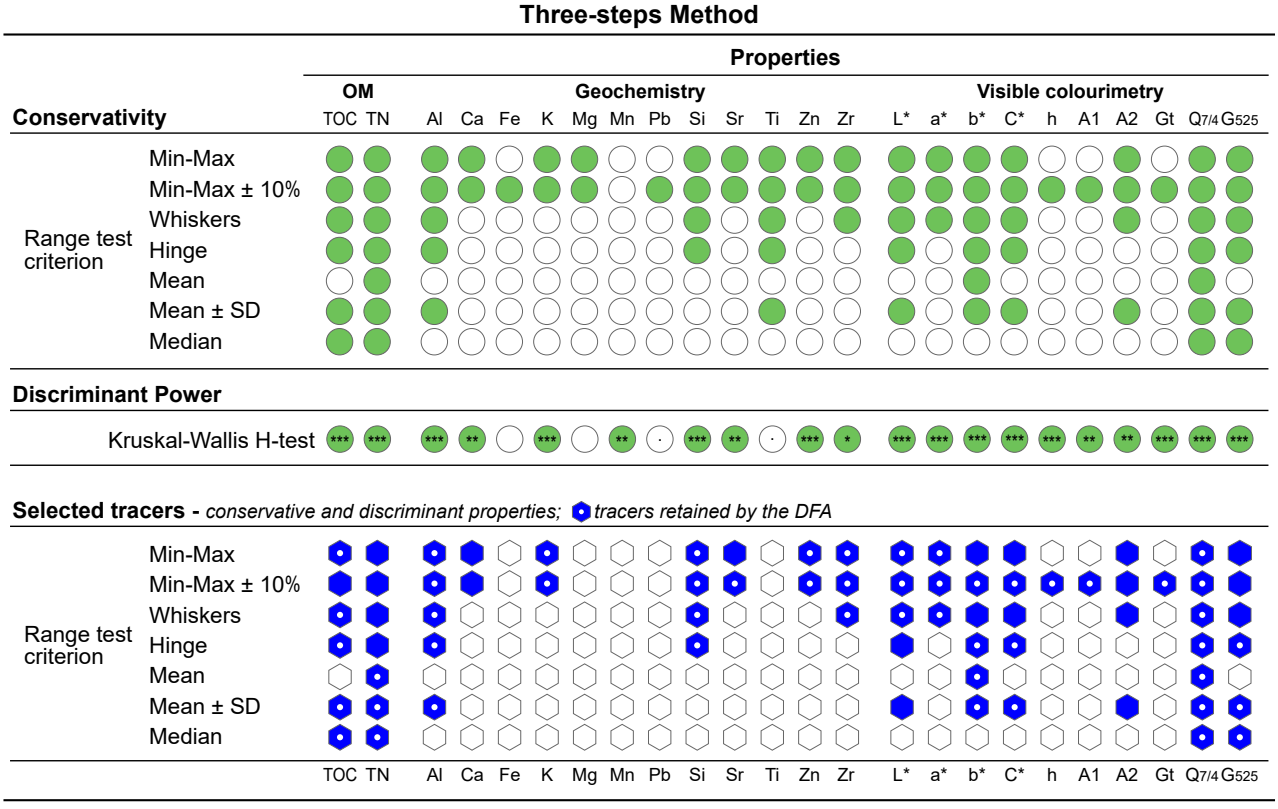

Figure 2. Tracer selection using the Three-Step Method according to different range test criteria. Green filled circles indicate a property that passed the conservativity (i.e. range test) or the discriminant power (Kruskal-Wallis H-test) tests. Blue filled diamonds indicate selected tracers (i.e. properties that passed both conservativity and discriminant power tests) and blue diamonds with white points indicate tracers retained by the discriminant function analysis (DFA) forward stepwise selection based on Wilk's lambda criterion. Kruskal-Wallis H-test's p-value significance: "***" p-value < 0.001; "**" p-value < 0.01; "*" p-value < 0.05; "." p-value < 0.1; " " p-value < 1. OM: organic matter; Min-Max: minimum-maximum; SD: standard deviation.

330 **Consensus method.**

The CI values indicate that TOC, TN, b* and C* were identified as conservative properties (Fig. 3). For these four properties, all sediment samples (n = 32) obtained a CR score above the threshold (i.e. 70), and were therefore considered relevant and kept in the list of tracers for further analysis.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

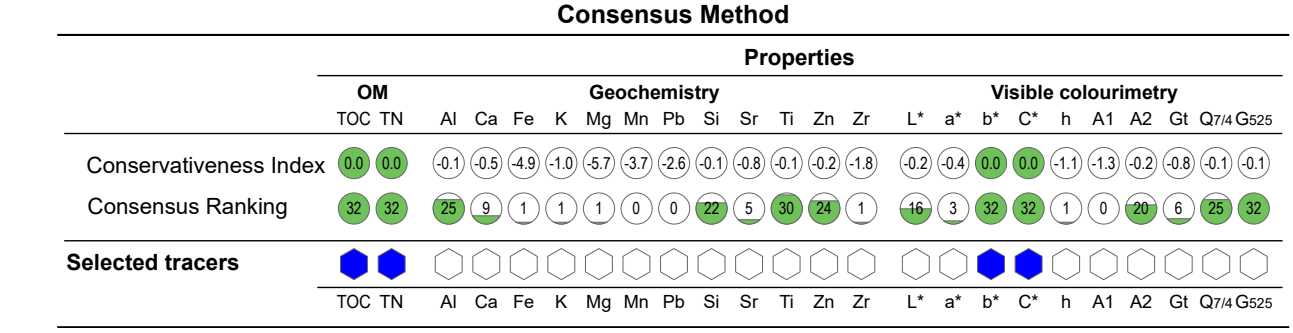

Figure 3. Tracer selection using the Consensus Method. Green filled circles indicate properties selected by the Conservativeness Index (i.e. CI = 0.0). The value in the circle refers to the number of sediment samples (out of 32) that obtained a Consensus Raking score above the threshold (i.e. CR $\geq$ 70). Blue filled diamonds indicate selected tracers. OM: organic matter.

335 Table 3 summarises the list of tracers selected by each approach.

Table 3. Tracers selected by the three-step method according to the range test criteria and the consensus method. Tracers underlined and written in bold were retained by the discriminant function analysis (DFA) forward stepwise selection based on Wilk's lambda criterion. Min-Max = minimum-maximum

Tracer selection method	Range test	**Selected tracers**
Three-step method	Min-Max	**TOC**, TN, **Al**, Ca, **K**, **Si**, Sr, **Zn**, **Zr**, **L***, **a***, b*, C*, A2, **Q7/4**, G_{525}
	Min-Max ± 10%	TOC, TN, **Al**, Ca, **K**, **Si**, **Sr**, **Zn**, **Zr**, **L***, **a***, **b***, **C***, **h**, **A1**, A2, **Gt**, **Q7/4**, G_{525}
	Whiskers	**TOC**, TN, **Al**, **Si**, **Zr**, **L***, **a***, b*, C*, A2, **Q7/4**, G_{525}
	Hinge	**TOC**, TN, **Al**, **Si**, **L***, **b***, **C***, **Q7/4**, **G_{525}**
	Mean	**TN**, **b***, **Q7/4**
	Mean ± SD	**TOC**, **TN**, **Al**, **L***, **b***, **C***, A2, **Q7/4**, G_{525}
	Median	**TOC**, **TN**, **Q7/4**, **G_{525}**
Consensus method		TOC, TN, b*, C*

3.2 Prediction accuracy of tracer selection methods

The prediction accuracy of the mean range test criterion and consensus method tracer selections were lower for each source compared to that obtained with the other three-step method range test criteria (Fig. 4). However, the difference in prediction accuracy was greater between sources within each tracer selection method than between methods. The effect of DFA stepwise
340 selection was moderate, mainly modifying the prediction accuracy of cropland and forest. On average, W50* values (+ 20%),

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

r^2 (+ 0.01) and NSE (+ 0.04) increased, while ME (- 37%), RMSE (- 8%) and CRPS (- 1) decreased. In short, uncertainty increased (W50*), bias decreased (ME) and performance slightly increased (r^2, NSE, CRPS*).

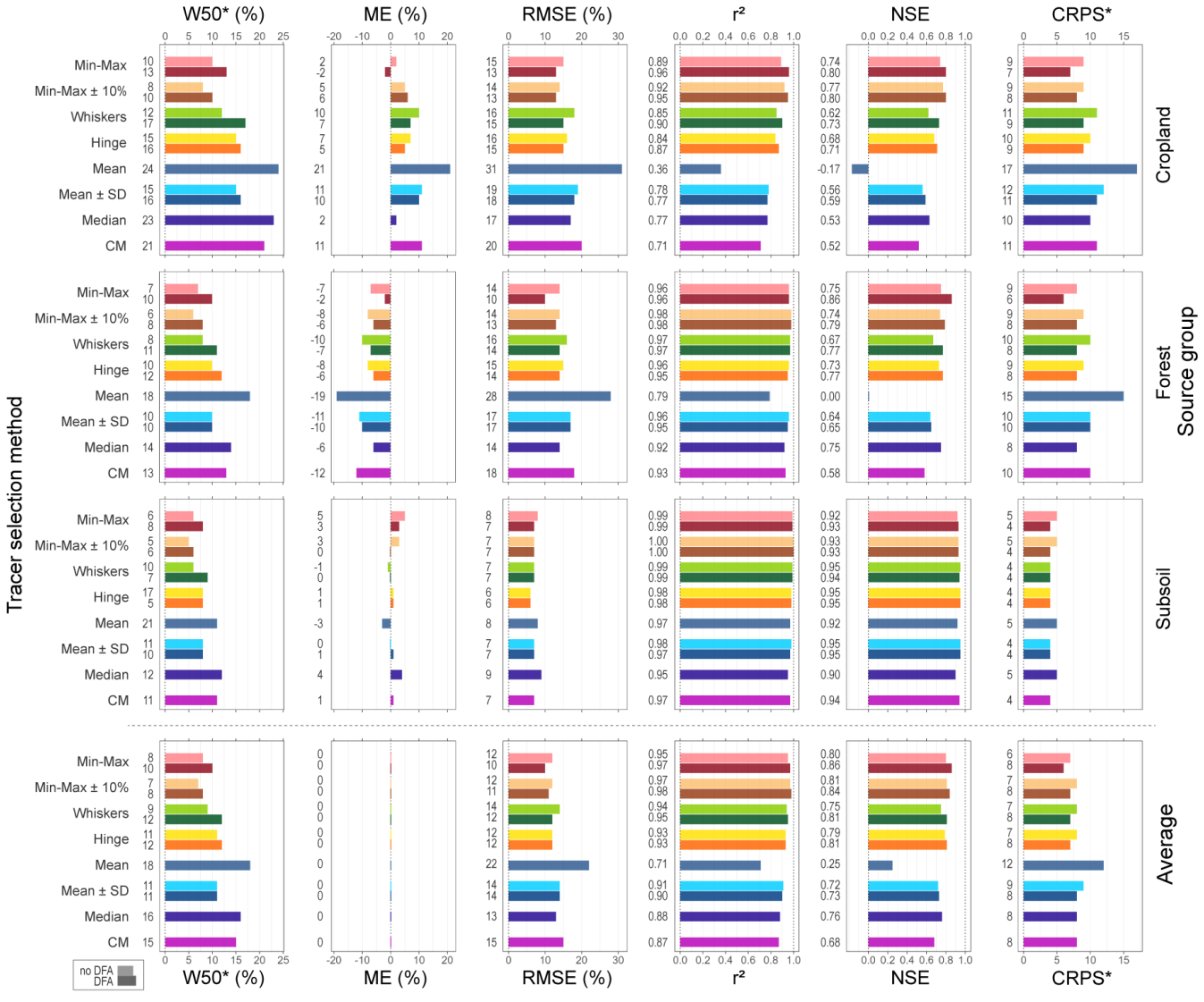

Figure 4. Summary of MixSIAR prediction statistics calculated on virtual mixtures (n = 138) for each tracer selection method: Three-Step Method range test criteria and Consensus Method for each individual soil source and average of all three sources for each method. Tracers selected by each method are listed in Table 3. ME: mean error; RMSE: root mean square error; NSE: Nash-Sutcliffe modelling efficiency coefficient; CRPS: continuous ranked probability score. Note: **Mean values per source.*

Regardless of the tracer selection method, the W50* values indicate a higher uncertainty for cropland, followed by forest and subsoil (Fig. 4). The consensus method, mean and median range test criteria showed higher W50* for each source (21-25,

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

345 13-18 and 11-12% for cropland, forest and subsoil, respectively) among the tracer selection methods. The lowest W50* values were obtained by minimum-maximum ± 10 % and minimum-maximum with and without DFA stepwise selection. Overall, minimum-maximum and minimum-maximum ± 10 % showed homogeneous values over the 0-100% contribution range for all sources (Fig. 5). The subsoil W50 curves showed no trend over the contribution range for most of the range test criteria, while there was a slight reduction in W50 values for the median criterion and the consensus method. However, the W50 value
350 for cropland and forest tended to increase for most of the range test criteria (i.e. whiskers, hinge, mean, mean ± SD, median) and the consensus method. This increase was more pronounced for the forest and also for the consensus method, the mean and median criteria. In addition, the W50 values were more scattered for these tree tracer selections, especially for the contributions below 60-70%. As observed for the W50* (Fig. 4), the DFA stepwise selection was associated with an increase in W50 values for all range test criteria.

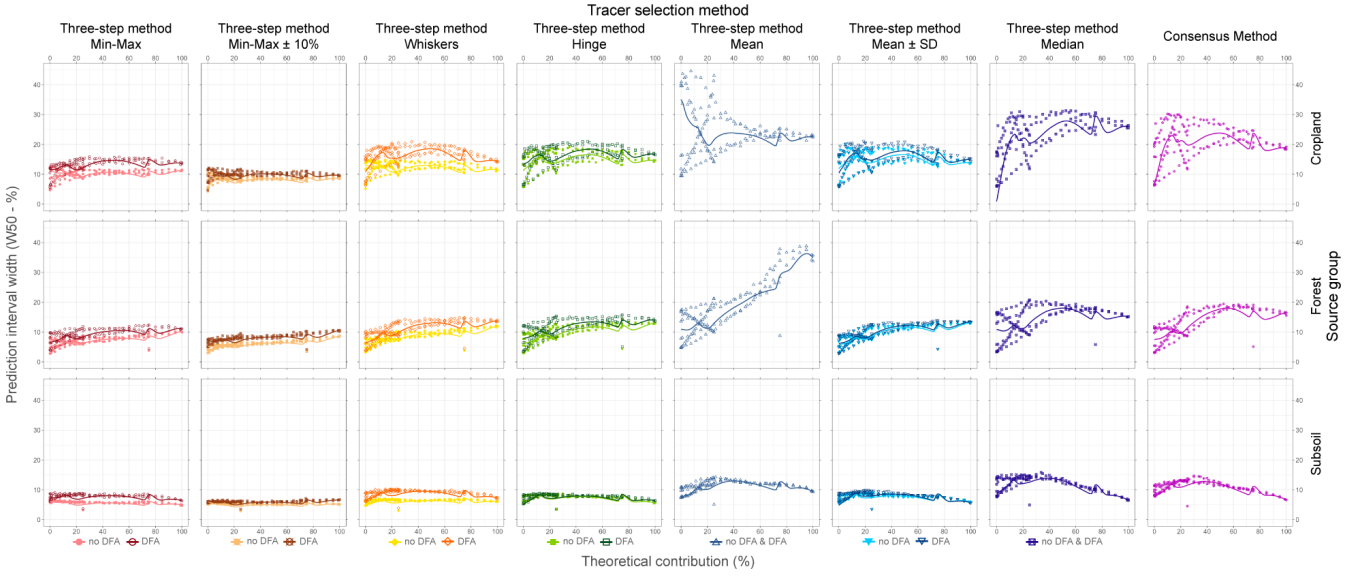

Figure 5. Relationship between virtual mixture source contributions and prediction interval width (W50) across sources: cropland, forest, subsoil, according to each tracer selection approach: three-step method range test criteria: minimum-maximum (red circle), minimum-maximum ± 10% (brown crossed circle) , whiskers (yellow/orange diamond), hinge (green square), mean (blue triangle), mean ± SD (blue triangle point down), median (blue crossed square), filled and empty symbols correspond to the selection before DFA (no DFA) and after DFA respectively, and consensus method (CM; purple circle plus).

355 Regarding the model residuals for all tracer selections except that obtained with the mean criteria, the error for subsoil was low (RMSE = 6-8%) with a small bias (ME = -3 - 5%). In contrast, forest and cropland errors were about twice higher (RMSE = 10-20%), with a negative bias for forest (ME = -19 - -2%) and a positive bias for cropland (ME = 2-11%). Accordingly, forest and cropland contributions were under- and over-predicted, respectively.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Then, in terms of model performance, the subsoil predictions were highly linear (r^2 = 0.97-0.99) and well predicted (NSE = 0.92-0.95). For the three-step method criteria, linearity and prediction quality were slightly improved in most cases. Nevertheless, a slight decrease in precision and accuracy was observed for contributions below 10% and above 80% (Fig. 6). Forest predictions were highly linear for most tracer selections (r^2 = 0.92-0.98), although their prediction quality was lower (NSE = 0.58-0.86) due to an under-estimation (i.e. ME values, Fig 6). This under-estimation was also confirmed by the CRPS values, which increased strongly and linearly above 20% of forest contribution (Fig. 6). Cropland predictions were relatively linear (r^2 = 0.71-0.89) but not well predicted (NSE = 0.52-0.80), representing the lowest performance among the sources. The CRPS curves were U-shaped, with a decrease in accuracy and precision for contributions below 60% and above 70% (Fig. 6). This pattern was also observed on the observed versus predicted plots in Fig. 7, with a tipping point at around 40 to 60% of the contribution. In addition, for a contribution below 60%, two types of behaviour were observed regardless of the tracer selection method considered (Fig. 6). One group of virtual mixtures was well predicted with a small bias and low CRPS values ($\leq$ 0.05), while another group was significantly less well predicted with a strong positive bias and higher CRPS values (Fig. 7). These two groups correspond to virtual mixtures with a dominant proportion of subsoil and forest, respectively. For contributions above 60%, these two groups converge and no difference can be observed. This suggests that selected tracers were able to successfully discriminate between subsoil and forest sources, while cropland was confused with one of the other two sources.

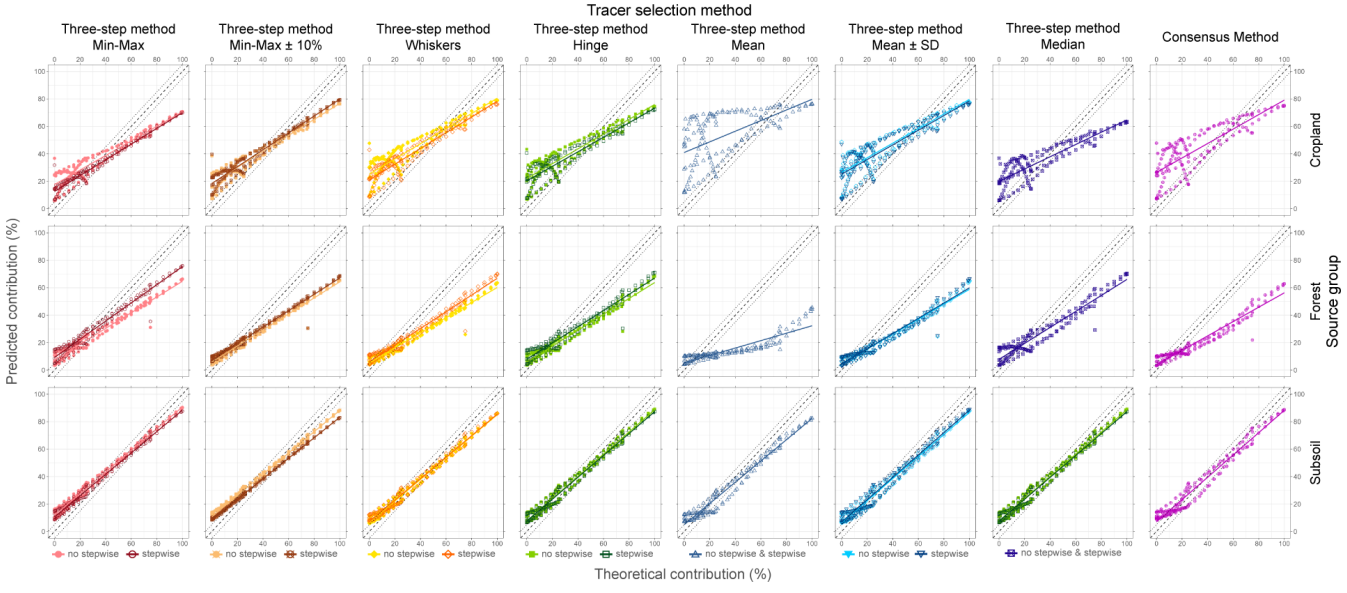

Figure 6. Relationship between predicted and observed source contributions: cropland, forest, subsoil, according to each tracer selection approach: three-step method range test criteria: minimum-maximum (red circle), minimum-maximum ± 10% (brown crossed circle) , whiskers (yellow/orange diamond), hinge (green square), mean (blue triangle), mean ± SD (blue triangle point down), median (blue crossed square), filled and empty symbols correspond to the selection before DFA (no DFA) and after DFA respectively, and consensus method (CM; purple circle plus).

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

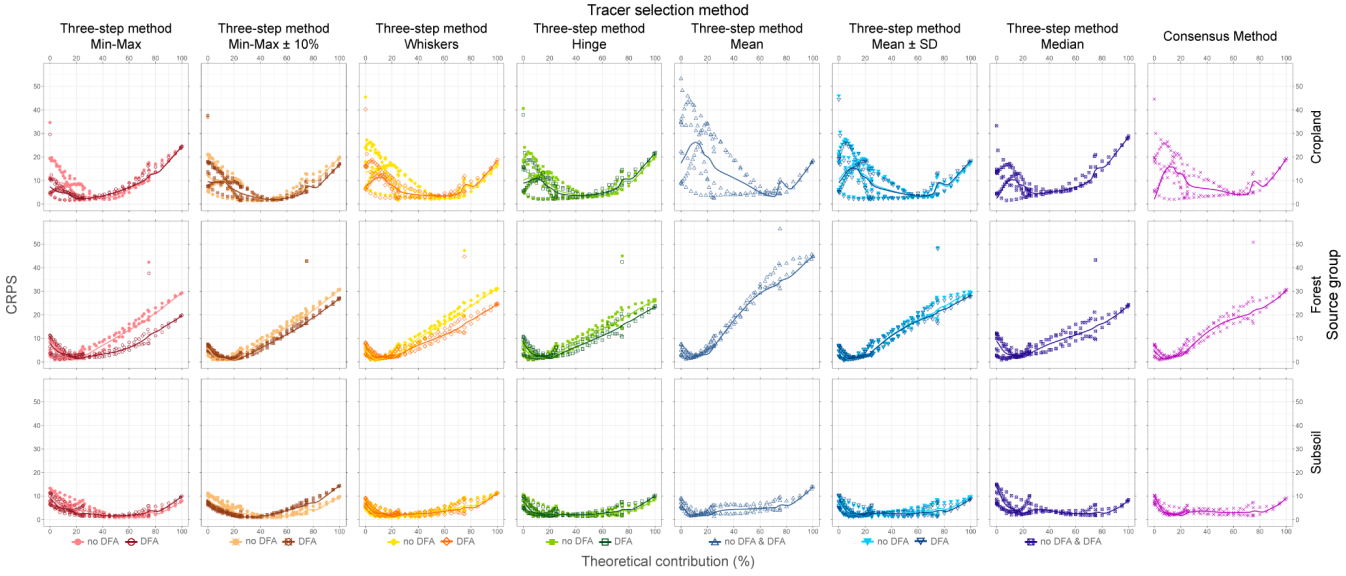

Figure 7. Relationship between virtual mixture source contributions and continuous ranked probability score (CRPS) across sources: three-step method range test criteria: minimum-maximum (red circle), minimum-maximum ± 10% (brown crossed circle) , whiskers (yellow/o-range diamond), hinge (green square), mean (blue triangle), mean ± SD (blue triangle point down), median (blue crossed square), filled and empty symbols correspond to the selection before DFA (no DFA) and after DFA respectively, and consensus method (CM; purple circle plus).

Among the selection of tracers, the mean range test criterion was the one that resulted in the lower prediction quality. Despite the fact that subsoil had a prediction quality close to that of the other methods, cropland and forest were poorly predicted (NSE = -0.17 and 0.00 respectively), modelled contributions being quite far from the theoretical values (Fig 6; 7)).

In addition, when looking at the predicted source contributions for the sediment core samples (i.e. 32 samples), according to the tracer selection method and the source, some sample predicted contributions were outside of the range of predicted contributions on the virtual mixtures. Among the sources, subsoil contributions were most frequently predicted outside of this range, with only 0 to 44% of the samples being within the respective virtual mixture contribution range. Only for the mean and mean ± SD without DFA stepwise criteria, 81% of the sediment samples remained within their respective range. For the hinge, mean, mean ± SD and median, the totality of sample cropland contributions corresponded to the respective virtual mixture contribution range and, for forest, the number of samples within correspond range was from 84 to 91%. The numbers of sample with predicted contributions for forest and subsoil remaining within similar ranges were much lower for minimum-maximum, minimum-maximum ± 10% and whiskers without DFA (0-41%) while they were higher for cropland (84-97%). However, for the minimum-maximum and minimum-maximum ± 10% with DFA, no sediment sample predicted contributions were within the range of virtual mixture predicted contributions for all sources. Overall, the DFA stepwise selection tended to reduce the number of matching sediment sample predicted contributions, regardless of the range test criteria. Regarding the consensus

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

method, 100% of forest, 31% of cropland and 19% of subsoil predicted contributions remained within the respective virtual mixture predicted contribution range.

3.3 Source contribution predictions

The tracer selection methods resulted in three types of source contribution trends: strong dominance of forest, strong dominance of cropland, and equivalent contribution of forest and cropland (Fig. 8).

For all approaches, subsoil contributions were low and stable, for most methods from 0 ± 1 to 8 ± 3 %. The mean and mean $\pm$ SD criteria without DFA led to higher and more variable subsoil contributions than other methods, 12 ± 5 % versus about 4 ± 2 %. The relationship between cropland and forest was different depending on the tracer selection method. For the consensus method, the mean and median criteria led to model outputs showing that cropland was dominant compared to forest along the sediment core. In contrast, according to the mean $\pm$ SD and hinge criteria, the contributions of cropland and forest were modelled to be similar or with only a slight dominance of forest. Finally, for the minimum-maximum, minimum-maximum $\pm$ 10% and whiskers criteria, the model calculated a strong dominance of forest over cropland. The DFA did not really affect the trends in the contributions of cropland and forest for the mean $\pm$ SD and hinge criteria, although it resulted in a significant smoothing of the contribution values for the minimum-maximum, minimum-maximum $\pm$ 10% and whiskers criteria. The contribution of cropland was greatly reduced from 16-27% to 2-8%, and the dominance of forest increased from 67-80% to 89-97%, resulting in an almost unique and stable contribution from forests along the entire core. However, the contributions of cropland and forest showed similar variations and tendencies along the sediment core for all tracer selection methods, although their values were significantly different. Samples taken at 9/10-10/11, 11/12-13/14, 19/20-20/21 and 24/25-07/28 cm depths were associated with higher contributions from forest and lower contributions from cropland compared to upper/lower samples.

The selected tracers did not provide the same intrinsic information and therefore did not provide the same ability to discriminate between the sources (Fig. 9). For most of the tracers, the following relationship between source group signatures (i.e. mean and sd of log transformed values) was observed: cropland showed intermediate values, with either forest (TOC, TN) or subsoil (Al, Si, L*, a*, b*, C*, A2, Q7/4, G_{525}) being the highest values. For most of them, cropland had similar values as those of forest (TOC, TN, Si, a*, b*, C*, A2, Q7/4) or subsoil (G_{525}). For a few tracers, cropland had the higher values with either subsoil (Zn, Gt) or forest (h) corresponding to the the lower values. However, for h, the signatures of cropland and subsoil were similar, and this similarity was also observed for Zn in cropland and forest, and for Gt in forest and subsoil sources. In addition, for some selected tracers (Ca, K, Sr, Zr, A1), all source signatures were all very close and hardly distinguishable, making the information derived from the relationships between sources less clear.

Depending on the tracer, sediment sample values did not show the same position in relation to the source signatures (Fig. 9. For some tracers, sample values were close to those of cropland (i.e. Zn, b*, A1, Q7/4), forest (Al, K, Zr, L*) or subsoil (a*, C*, h). However, for most of the tracers (i.e. TOC, TN, Ca, Si, Sr, A2, Gt, G_{525}) samples values were intermediate between those of cropland and forest.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

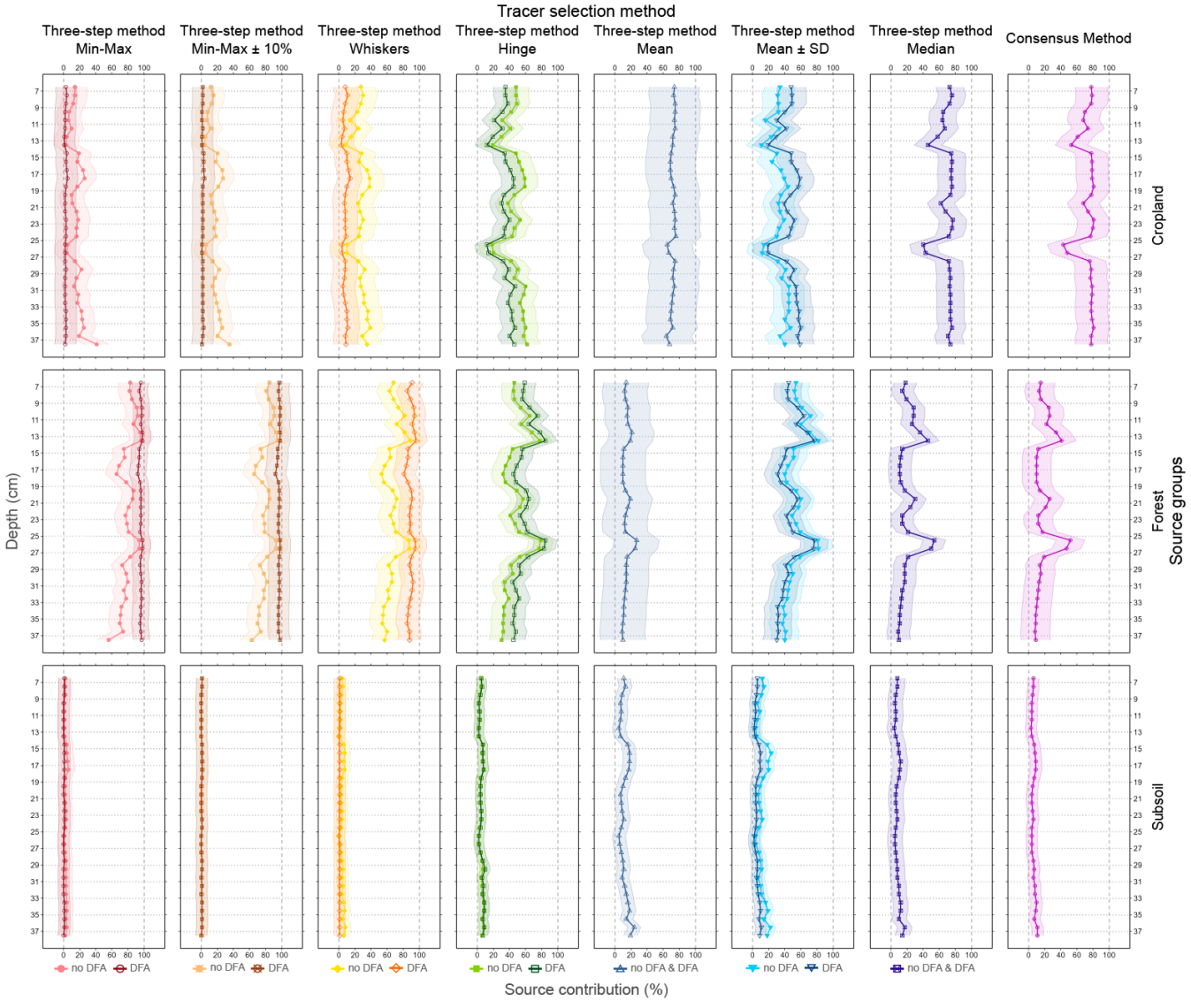

Figure 8. Predicted source contributions for the sediment core samples according to each tracer selection approach: the different range tests criteria of the three-step method: minimum-maximum (red circle), minimum-maximum ± 10% (brown crossed circle), whiskers (yellow/orange diamond), hinge (green square), mean (blue triangle), mean ± SD (light blue triangle point down), median (purple crossed square), empty and filled symbols correspond to the use of DFA or not, and consensus method (CM; light purple circle plus). The error buffer ribbons around the plotted values correspond to the respective RMSE values calculated on virtual mixtures.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

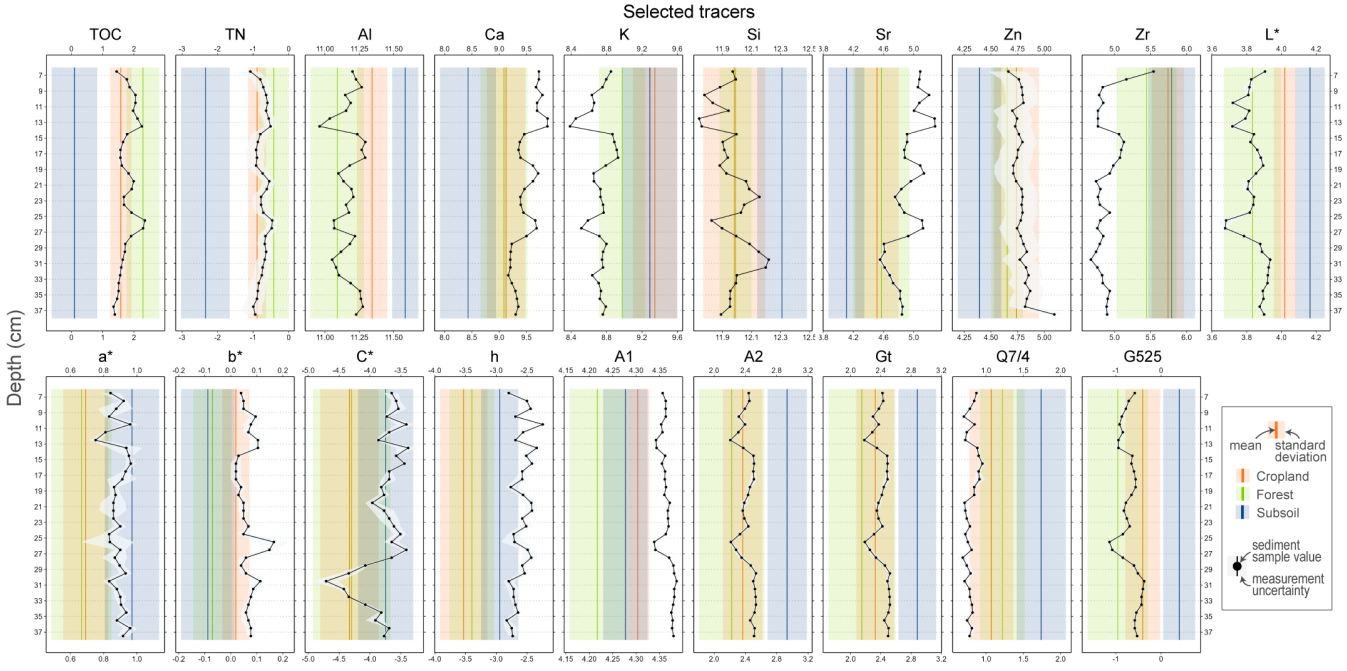

Figure 9. Log transformed tracer values in sediment core samples (black dot, and measurement uncertainty = grey buffer) and potential source signatures (vertical line = mean value; buffer zone = standard deviation).

4 Discussion

4.1 Conservativity assessment

The different tests used to assess conservativeness, the three-step method range test using different criteria and the consensus method conservativeness index, resulted in a more or less restrictive selection of properties. These tests can be divided into three groups according to the number of properties selected, from the least restrictive (minimum-maximum, minimum-maximum $\pm$ 10%, whiskers), the moderately restrictive (hinge, mean $\pm$ SD), to the more restrictive (conservativeness index, mean, median). Overall, the conservativity tests tended to mainly identify the same properties as being conservative: TOC, TN, b*, C*, and Q7/4.

In this study, all the tests identified organic matter properties as conservative (i.e. TOC and TN), except for the mean criterion, which did not select TOC. To assess the composition of organic matter, whether from terrestrial and/or freshwater-originating material, the distribution of δ^{13}C versus C/N ratio can be compared to thresholds (Lamb et al., 2006). In a previous study conducted on a sediment core sampled from the same site in Hayama Lake in 2014, Huon et al. (2018) concluded to the absence of an autochtonous input of freshwater-originating organic matter, which was confirmed in our data (see Supplementary material fig. S1). Organic matter, and especially carbon are expected to provide a conservative property (García-Comendador et al., 2023) and this is supported by their widespread selection.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Among the geochemical properties, Al, Ti and Si were frequently selected by range tests. This can be explained by their well-known low solubility (Meybeck and Helmer, 1989; Phillips and Greenway, 1998; Koiter et al., 2013), and more specifically for Al and Si, which are constituents of clay sheets, and given the granitic geological context of the study area (see Supplementary Material Fig. S2). However, these geochemical properties are sensitive to grain size sorting occurring along the transport pathway as they are clays sheets constituents. Other geochemical properties that were rarely selected (i.e. Fe, Pb, Ca, K) or systematically rejected, such as Mn, are associated with higher solubility or higher desorption susceptibility (Koiter et al., 2013; García-Comendador et al., 2023; Meybeck and Helmer, 1989; Phillips and Greenway, 1998).

Among the visible colour indices, all conservativity tests selected Q7/4, while other indices such as G_{525}, b*, C*, and L* were widely selected. The colour indices L*, b* and C* were highly correlated (Pearson r^2 from -0.61 to -0.77, see Supplementary Materials Table. II) with organic matter properties (i.e. TOC and TN) that were also identified as conservative. This is consistent with the fact that organic matter and iron oxides are the main soil colouring elements. The existence of different visible colour spaces allows to selected the most appropriate space and indices according to the local context (Viscarra Rossel et al., 2006). These colour spaces are interconnected and transformation of indices can be achieved with simple mathematical formula. However, care must be taken to avoid multicolinerarity when using multiple indices from different colour spaces together, as redundancy of information tends to degrade modelling accuracy (Cox et al., 2023). However, the visible colour parameters are quite sensitive to spatial and temporal variations (García-Comendador et al., 2023). The acquisition of other spectral regions such as Vis-NIR, NIR and MIR appears to be more robust (Chen et al., 2023), especially as these regions are a powerful and reliable way to obtain extensive range of properties and information on the sample with the advantages of being rapid, cost-effective and non-destructive measurement (Soriano-Disla et al., 2014). Of note, in order to ensure the exchange of spectra within the community, the adoption of a common protocol and/or the provision of calibration spectra using inter-calibration samples should be discussed, as spectra tend to be instrument-dependent (Pimstein et al., 2011).

Regarding the conservativeness index, it was a restrictive method compared to most range test criteria, selecting only four properties, which is in the same order of magnitude as the number of properties selected by the mean and median range test criteria. Nevertheless, some properties selected by other range test criteria, which were a priori conservative properties, showed a score close to the threshold of 0.00 defined by Lizaga et al. (2020a). Thus, Al, Si, Ti, $Q_{7/4}$ and G_{525} obtained a conservativeness index equal to -0.1, while it was equal to -0.2 for Zn, L* and A2. This could indicate that properties that yielded a conservativeness index score close to the threshold (i.e. 0.00) were not strictly conservative (e.g. size sorting for Al, Si and Ti). We suggest that either the conservativeness index can be used to identify less conservative properties in a selection resulting from different tests.

Of all the different tests compared to assess conservativeness, the mean $\pm$ SD and hinge range test criteria selected the most relevant sets of properties. The mean and median range test criteria and the conservativeness index also provided relevant sets of properties, but they selected too few tracers, which could be a limitation for fingerprinting modelling.

As this study shows, a priori and field knowledge are essential to assess the relevance of conservativeness assessments (Laceby et al., 2015; Koiter et al., 2018; Batista et al., 2019). However, this knowledge is not sufficient and not usable for all measured properties, especially for relatively new properties as colour parameters, due to their complex relationship with

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

environmental processes. Other studies developed particle size and organic matter corrections, which were shown to be effective (Laceby et al., 2017; Koiter et al., 2018). However, they require additional measurements and are site-specific (Koiter et al., 2018), which is not in line with the desired simplification of the technique. Therefore, it is essential to develop cost-, time-effective and generalised methods (Koiter et al., 2018).

4.2 Tracer selection and contribution modelling

The consensus ranking index was designed to select properties based on their relevance to prediction, as they maximise convergence. Of the 24 properties considered in the current research, only the TOC, TN, b*, L*, and G_{525} reached a consensus ranking score above 70 (Fig. 3). These properties were redundant, as they showed very similar trends along the sediment core (Fig. 9). By maximising the consensual aspect of the properties, consensus raking favours redundant properties and discards descendant properties that could nevertheless be relevant, and this may lead to an information gap. Descendant properties can provide complementary information, as properties do not discriminate between sources in the same way.

Indeed, the selection of consistent properties can result in the prediction of the dominant source while ignoring other sources. In this study, the consensus ranking selected properties for which sediment values were close to cropland signature or were intermediate between cropland and forest values (Fig. 9), which explains the dominance of cropland contribution calculated by CM (Fig. 8).

Conversely, the Discriminant Function Analysis (DFA) stepwise procedure of the three-step method aims to maximise the difference between properties to minimise redundant information, which could reduce the prediction quality of MixSIAR (Cox et al., 2023). In this study, the use of the DFA improved model accuracy for all the range test criteria (Fig. 4), but it also decreased the concordance between results obtained for virtual mixtures and sediment sample contributions. In each case, the number of sediment for which predicted contributions fell outside of the range of the virtual mixture predicted contributions increased. In addition, DFA application tended to reduce source contribution variations (Fig. 8), which is not in line with the sediment fingerprinting objectives. In the case of a reliable tracer selection, as offered by the mean $\pm$ SD and hinge range, DFA application is less relevant. Overall, the impact of the DFA on tracer selection modelling outputs needs to be clarified. It is therefore preferable to run the model with both selections, and keep the most relevant one.

From the different tracer selections, which were all distinct, three main tendencies were observed in terms of modelled contributions (Fig. 8). On the one hand, extensive tracer selections, such as minimum-maximum, minimum-maximum $\pm$ 10% and whiskers criteria, resulted in a modelled dominance of the forest contribution over cropland and subsoil. On the other hand, restrictive selections of tracers, such as the consensus method, mean and median criteria, resulted in a modelled dominant contribution of cropland. Finally, methods that selected an intermediate number of tracers, i.e. mean $\pm$ SD and median criteria, led to the prediction of a balanced contribution between forest and cropland. This major impact of the tracer selection methods on the source contribution modelling was demonstrated by multiple studies (Laceby et al., 2015; Palazón et al., 2015; Smith et al., 2018; Gaspar et al., 2019).

However, in the current research, all methods agreed on a low subsoil contribution associated with high modelling accuracy statistics (Fig. 4). It can be assumed that since all tracer selection methods selected TN, b* and C*, they allowed to predict

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

the general trend of the subsoil contribution behaviour. Thus, methods that selected additional tracers, such as mean ± SD and median criteria, did not result in significant modifications of the trend. However, the methods that selected a large number of tracers (i.e. minimum-maximum, minimum-maximum ± 10%, whiskers) resulted in a large reduction or even disappearance of the subsoil contribution. The need for few tracers to predict subsoil with a good prediction accuracy can be explained by its significantly different signature compared to forest and cropland (Fig. 9). Most of the modelling limitations were related to the prediction quality for cropland and forest, as their respective signatures were very close to each other for many tracers (Fig. 9). However, the same variations, i.e. higher forest and lower cropland contributions, were observed for most methods. These variations could be explained by the information provided through the incorporation of TN and TOC contents, which were the most frequently selected tracers. That is consistent with the results of Huon et al. (2018), which associated a higher TOC content with a high forest contribution. Furthermore, for the methods that selected Al, these variations were sharper and more detailed, especially for samples collected from the upper part of the sediment core (i.e. depth of 7 to 17 cm).

Of note, additional metrics, such as sensitivity analysis or variable importance approach, could provide a more detailed understanding of the role of each tracer in contribution predictions (Russi et al., 2008; Bennett et al., 2013; Wei et al., 2015).

4.3 Assessing modelling prediction accuracy

The generation of virtual mixtures allowed the evaluation of model prediction accuracy for a wide range of source contributions. The use of several metrics allowed to describe different aspects of the modelling (i.e. residuals, accuracy and precision) and to better interpret the prediction on real sediment samples (Latorre et al., 2021; Batista et al., 2022). The graphical representation of the metrics (Figs. 5, 6, and 7) allowed to identify ranges of source contributions with different prediction accuracies. This understanding of the model supports a better appreciation and interpretation of predictions on sediment samples.

However, for most of the tracer selections studied here, some or all of the predicted contributions of sediment samples fell outside of the range of the virtual mixture predicted contributions (Fig. 6 and 8). It should be noted that the range of virtual mixture predicted contributions defines the minimum and maximum possible predictions for a mixture, virtual or real, that may be obtained with a given set of tracers. Therefore, the occurrence of sediment samples with predicted contributions outside of this range implies that tracer values were different from those expected when generating virtual mixtures. In fact, when generated as a simple proportional mixture of source tracer properties, virtual mixtures are based on the assumption of a strict conservativity of properties (Palazón et al., 2015; Smith et al., 2018; Batista et al., 2022). Therefore, the use of non-conservative tracers will result into the generation of virtual mixture property values that differ from those observed in reality. As a result, the evaluation of model prediction accuracy with these biased virtual mixtures provides an unrealistic evaluation of the model (Batista et al., 2022). This was observed here with extensive tracer selections, such as minimum-maximum, minimum-maximum ± 10% and whiskers, that led to better prediction accuracy metrics than more restrictive ones (i.e. hinge, mean, mean ± SD, median, consensus method) (Fig. 4). Nevertheless, these extensive selections resulted in a much lower number, and sometimes the absence of sediment sample predicted contributions matching those of virtual mixtures. This reduces confidence in the significance of the calculated metrics. It is therefore essential to correctly identify conservative properties in order to generate realistic virtual mixtures and thus allow a correct evaluation of the model prediction accuracy.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

We recommend to assess the number of sample predicted contributions which are within the virtual mixtures range of predicted contributions (i.e. minimum and maximum) to assess possible divergence of properties and therefore the reliability of evaluation metrics calculated with virtual mixtures.

In addition, all potential tracing properties are likely not modified in the same way by erosion processes (Koiter et al., 2013), and assessing quantitatively the extent to which a property is affected (i.e. depletion or enrichment) (García-Comendador et al., 2023) could support the generation of more realistic virtual mixtures, and thus bring model evaluation closer to reality.

5 Conclusions

In this study, we compared two source sediment fingerprinting tracer selection methods, the most conventional three-step method and the consensus method, on their tracer selections, and their source contribution predictions for the same dataset. Conservativity tests of both methods were compared, along with several three-steps method range test criteria and the consensus method conservativeness index. The different conservativity tests resulted in a more or less restrictive identification of conservative properties. On the one hand, the minimum-maximum, minimum-maximum $\pm$ 10% and whiskers range test criteria were not restrictive enough and selected non-conservative properties. On the other hand, mean and median criteria and the conservativeness index were too restrictive and selected too few properties, which can lead to limitations when modelling source contributions in target samples. The mean $\pm$ SD and hinge criteria resulted in a moderate and relevant identification of conservative properties. In addition, within the three-step method, the impact of Discriminant Function Analysis (DFA) stepwise selection on tracer selection was evaluated. Although the use of DFA tracer selections showed an improvement of model accuracy, it was associated with a decrease of concordance between virtual mixtures and sediment sample contributions, and a smoothing of contribution variations along the sediment core. However, in order to clarify DFA impact on model outputs, we recommend to run and compare the impact of both selections in other studies.

Although the different methods resulted in different selections of tracers, three main contribution tendencies were observed in relation with the number of selected tracers: extensive selections of tracers resulted in a strong dominance of forest, restrictive selections of tracers in a dominant contribution of cropland, and a balanced contributions of forest and cropland was obtained when selecting an intermediate number of tracers.

To assess modelling accuracy, 138 virtual mixtures were generated as proportional mixtures of sources. Several modelling accuracy metrics were computed for each tracer selection. These metrics and their associated representations provided a better understanding of each tracer selection impact on modelling uncertainty. However, for most of the tracer selections, a great or lesser number of sample predicted contributions fell outside of the range of the virtual mixture predicted contributions. This implies that tracer selections contained non-conservative or not strictly conservative tracers. It is therefore fundamental to correctly identify conservative properties to avoid the generation of biased virtual mixtures and therefore the calculation of unrealistic modelling accuracy metrics.

Among the compared methods, the three-step method using the mean $\pm$ SD or hinge range test criteria selected the most reliable set of tracers. However, the high variability of selected tracers among the three-step method, the consensus method

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

and within the three-step method range test criteria, and their strong impact on modelling output results, may raise concerns regarding the validity of quantitative outputs, which may not meet the ultimate goal sought for by fingerprinting approaches as they are currently implemented by the scientific community. Consequently, to avoid a potential loss of confidence of stakeholders regarding the validity of the outputs of this method in the future, it is essential to take as much care as possible to conduct an accurate and reliable identification of conservative properties, as the whole methodology and the ultimate results rely on this initial step. This is fundamental both for improving our understanding of erosion and sedimentation processes and for guiding the implementation of effective landscape management measures. Accordingly, we encourage our colleagues from the scientific community to share their tracing datasets obtained in contrasted environmental conditions around the world in order to contribution to the further improvement and evaluation of sediment source fingerprinting techniques.

Code and data availability. The dataset is available on line at Chalaux-Clergue et al. (2022). The code for running models, summarizing and plotting results is available in the supplementary material. In order to facilitate the implementation of the presented framework, an R package: fingR (Chalaux-Clergue and Bizeul, 2023), with all the functions used in the current study, has been developed and is freely accessible.

Author contributions. TCC, RB, OE and PB contributed to the conceptualisation of the study. TCC and RB developed the analysis and wrote the manuscript. TCC wrote the code and performed the modelling. OE, NMC, PB, and JPL revised the manuscript and contributed to the text. OE and JPL sampled source soils, OE supervised the laboratory analyses and TCC conducted the laboratory analyses on the sediment core samples.

Competing interests. Olivier Evrard is a member of the editorial board of the journal.

Acknowledgements. The collection and the analysis of the soil and sediment samples were funded by the TOFU (ANR-11-JAPN-001) and the AMORAD (ANR-11-RSNR-0002) projects, under the supervision of the French National Research Agency (ANR, Agence Nationale de la Recherche). The support of CEA (Commissariat à l'Energie Atomique et aux Energies Alternatives, France), CNRS (Centre National de la Recherche Scientifique, France) and JSPS (Japan Society for the Promotion of Science) through the funding of PhD fellowships (H. Lepage, H. Jaegler, T. Chalaux-Clergue) and collaboration projects (grant no. PRC CNRS JSPS 2019-2020, no.10; CNRS International Research Project – IRP – MITATE Lab) is also gratefully acknowledged. T. Chalaux-Clergue obtained a JSPS grant to spend his second year of PhD at Kyoto Prefectural University (Oct. 2022 - Sept. 2023) (Grant number PE22708). This work was also supported by ERAN (Environmental Radioactivity Research Network Center) grants I-21-22 and I-22-24, and a mini-project from the FIRE (Fédération Ile-de-France de Recherche en Environnement - CNRS FR3020 FIRE, France) entitled RICOR ('Reconstructing the Impact of Changing land Occupation on the fate of Radionuclides').

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

References

Batista, P. V. G., Laceby, J. P., Silva, M. L. N., Tassinari, D., Bispo, D. F. A., Curi, N., Davies, J., and Quinton, J. N.: Using Pedological Knowledge to Improve Sediment Source Apportionment in Tropical Environments, Journal of Soils and Sediments, 19, 3274–3289, https://doi.org/10.1007/s11368-018-2199-5, 2019.

Batista, P. V. G., Laceby, J. P., and Evrard, O.: How to evaluate sediment fingerprinting source apportionments, Journal of Soils and Sediments, 22, 1315–1328, https://doi.org/https://doi.org/10.1007/s11368-022-03157-4, 2022.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Scientific data, 5, 1–12, https://doi.org/https://doi.org/10.1038/sdata.2018.214, 2018.

Bennett, N. D., Croke, B. F., Guariso, G., Guillaume, J. H., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T., Norton, J. P., Perrin, C., et al.: Characterising performance of environmental models, Environmental modelling & software, 40, 1–20, 2013.

Bing, H., Wu, Y., Liu, E., and Yang, X.: Assessment of heavy metal enrichment and its human impact in lacustrine sediments from four lakes in the mid-low reaches of the Yangtze River, China, Journal of Environmental Sciences, 25, 1300–1309, https://doi.org/10.1016/S1001-0742(12)60195-8, 2013.

Blake, W. H., Walsh, R. P. D., Sayer, A. M., and Bidin, K.: Quantifying Fine-Sediment Sources in Primary and Selectively Logged Rainforest Catchments Using Geochemical Tracers, Water, Air, & Soil Pollution: Focus, 6, 615–623, https://doi.org/10.1007/s11267-006-9046-1, 2006.

Brosinsky, A., Foerster, S., Segl, K., López-Tarazón, J. A., Piqué, G., and Bronstert, A.: Spectral fingerprinting: characterizing suspended sediment sources by the use of VNIR-SWIR spectral information, Journal of Soils and Sediments, 14, 1965–1981, https://doi.org/10.1007/s11368-014-0927-z, 2014.

Chalaux-Clergue, T. and Bizeul, R.: fingR: A framework for Sediment Source Fingerprinting, https://doi.org/https://doi.org/10.5281/zenodo.8293596, r package version 1.0.0, 2023.

Chalaux-Clergue, T., Evrard, O., Durand, R., Caumon, A., Hayashi, S., Tsuji, H., Huon, S., Vaury, V., Wakiyama, Y., Nakao, A., Laceby, J. P., and Onda, Y.: Organic matter, geochemical and colorimetric properties of potential source material, target sediment and laboratory mixtures for conducting sediment fingerprinting approaches in the Mano Dam Reservoir (Hayama Lake) catchment, Fukushima Prefecture, Japan., https://doi.org/10.5281/ZENODO.7081094, version Number: 1 Type: dataset, 2022.

Chartin, C., Evrard, O., Onda, Y., Patin, J., Lefèvre, I., Ottlé, C., Ayrault, S., Lepage, H., and Bonté, P.: Tracking the early dispersion of contaminated sediment along rivers draining the Fukushima radioactive pollution plume, Anthropocene, 1, 23–34, https://doi.org/10.1016/j.ancene.2013.07.001, 2013.

Chartin, C., Evrard, O., Laceby, J. P., Onda, Y., Ottlé, C., Lefèvre, I., and Cerdan, O.: The impact of typhoons on sediment connectivity: lessons learnt from contaminated coastal catchments of the Fukushima Prefecture (Japan), Earth Surface Processes and Landforms, 42, 306–317, https://doi.org/https://doi.org/10.1002/esp.4056, 2017.

Chen, D., Dai, W., Li, M., Wang, B., Zeng, Y., Ni, L., Fang, N., and Shi, Z.: Accuracy verification of optical fingerprinting methods in sediment tracing study, Hydrological Processes, 37, e14 870, https://doi.org/10.1002/hyp.14870, 2023.

Collins, A. and Walling, D.: Selecting fingerprint properties for discriminating potential suspended sediment sources in river basins, Journal of Hydrology, 261, 218–244, https://doi.org/10.1016/S0022-1694(02)00011-2, 2002.

Collins, A., Walling, D., and Leeks, G.: Source type ascription for fluvial suspended sediment based on a quantitative composite fingerprinting technique, Catena, 29, 1–27, https://doi.org/https://doi.org/10.1016/S0341-8162(96)00064-1, 1997a.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Collins, A., Walling, D., Webb, L., and King, P.: Apportioning catchment scale sediment sources using a modified composite fingerprinting technique incorporating property weightings and prior information, Geoderma, 155, 249–261, 2010.

Collins, A., Williams, L., Zhang, Y., Marius, M., Dungait, J., Smallman, D., Dixon, E., Stringfellow, A., Sear, D., Jones, J., and Naden, P.: Catchment source contributions to the sediment-bound organic matter degrading salmonid spawning gravels in a lowland river, southern England, Science of The Total Environment, 456-457, 181–195, https://doi.org/10.1016/j.scitotenv.2013.03.093, 2013.

Collins, A., Pulley, S., Foster, I., Gellis, A., Porto, P., and Horowitz, A.: Sediment source fingerprinting as an aid to catchment management: A review of the current state of knowledge and a methodological decision-tree for end-users, Journal of Environmental Management, 194, 86–108, https://doi.org/10.1016/j.jenvman.2016.09.075, 2017.

Collins, A. L. and Walling, D. E.: Documenting catchment suspended sediment sources: problems, approaches and prospects, Progress in Physical Geography: Earth and Environment, 28, https://doi.org/10.1191/0309133304pp409ra, 2004.

Collins, A. L., Walling, D. E., and Leeks, G. J.: Fingerprinting the origin of fluvial suspended sediment in larger river basins: combining assessment of spatial provenance and source type, Geografiska Annaler: Series A, Physical Geography, 79, 239–254, https://doi.org/https://doi.org/10.1111/j.0435-3676.1997.00020.x, 1997b.

Collins, A. L., Blackwell, M., Boeckx, P., Chivers, C.-A., Emelko, M., Evrard, O., Foster, I., Gellis, A., Gholami, H., Granger, S., et al.: Sediment source fingerprinting: benchmarking recent outputs, remaining challenges and emerging themes, Journal of Soils and Sediments, 20, 4160–4193, https://doi.org/https://doi.org/10.1007/s11368-020-02755-4, 2020.

Cox, T., Laceby, J. P., Roth, T., and Alewell, C.: Less is more? A novel method for identifying and evaluating non-informative tracers in sediment source mixing models, Journal of Soils and Sediments, 23, 3241–3261, https://doi.org/10.1007/s11368-023-03573-0, 2023.

Dabrin, A., Bégorre, C., Bretier, M., Dugué, V., Masson, M., Le Bescond, C., Le Coz, J., and Coquery, M.: Reactivity of particulate element concentrations: apportionment assessment of suspended particulate matter sources in the Upper Rhône River, France, Journal of Soils and Sediments, 21, 1256–1274, https://doi.org/10.1007/s11368-020-02856-0, 2021.

Debnath, A., Singh, P. K., and Chandra Sharma, Y.: Metallic contamination of global river sediments and latest developments for their remediation, Journal of Environmental Management, 298, 113 378, https://doi.org/10.1016/j.jenvman.2021.113378, 2021.

Debret, M., Sebag, D., Desmet, M., Balsam, W., Copard, Y., Mourier, B., Susperrigui, A.-S., Arnaud, F., Bentaleb, I., Chapron, E., et al.: Spectrocolorimetric interpretation of sedimentary dynamics: The new "Q7/4 diagram", Earth-Science Reviews, 109, 1–19, https://doi.org/https://doi.org/10.1016/j.earscirev.2011.07.002, 2011.

Evrard, O., Laceby, J. P., Ficetola, G. F., Gielly, L., Huon, S., Lefèvre, I., Onda, Y., and Poulenard, J.: Environmental DNA provides information on sediment sources: A study in catchments affected by Fukushima radioactive fallout, Science of The Total Environment, 665, 873–881, https://doi.org/10.1016/j.scitotenv.2019.02.191, 2019.

Evrard, O., Chaboche, P.-A., Ramon, R., Foucher, A., and Laceby, J. P.: A global review of sediment source fingerprinting research incorporating fallout radiocesium (137Cs), Geomorphology, 362, 107 103, https://doi.org/10.1016/j.geomorph.2020.107103, 2020a.

Evrard, O., Durand, R., Nakao, A., Laceby, P. J., Lefèvre, I., Wakiyama, Y., Hayashi, S., Asanuma-Brice, C., and Cerdan, O.: Impact of the 2019 typhoons on sediment source contributions and radiocesium concentrations in rivers draining the Fukushima radioactive plume, Japan, Comptes Rendus. Géoscience, 352, https://doi.org/10.5802/crgeos.42, 2020b.

Evrard, O., Batista, P. V., Dabrin, A., Foucher, A., Frankl, A., García-Comendador, J., Huguet, A., Lake, N., Lizaga, I., Martínez-Carreras, N., et al.: Improving the design and implementation of sediment fingerprinting studies: Summary and outcomes of the TRACING 2021 Scientific School, Journal of soils and sediments, pp. 1–14, https://doi.org/https://doi.org/10.1007/s11368-022-03203-1, 2022.

FAO: Soil erosion: the greatest challenge for sustainable soil management, FAO, p. 100, https://doi.org/20.500.12592/szbxmk, 2019.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Farias Amorim, F., Jacques Agra Bezerra da Silva, Y., Cabral Nascimento, R., Jacques Agra Bezerra da Silva, Y., Tiecher, T., Williams Araújo do Nascimento, C., Paolo Gomes Minella, J., Zhang, Y., Ram Upadhayay, H., Pulley, S., and Collins, A. L.: Sediment source apportionment using optical property composite signatures in a rural catchment, Brazil, CATENA, 202, 105 208, https://doi.org/10.1016/j.catena.2021.105208, 2021.

García-Comendador, J., Martínez-Carreras, N., Fortesa, J., Company, J., Borràs, A., Palacio, E., and Estrany, J.: In-channel alterations of soil properties used as tracers in sediment fingerprinting studies, CATENA, 225, 107 036, https://doi.org/10.1016/j.catena.2023.107036, 2023.

Gaspar, L., Blake, W. H., Smith, H. G., Lizaga, I., and Navas, A.: Testing the sensitivity of a multivariate mixing model using geochemical fingerprints with artificial mixtures, Geoderma, 337, 498–510, https://doi.org/https://doi.org/10.1016/j.geoderma.2018.10.005, 2019.

Gateuille, D., Owens, P. N., Petticrew, E. L., Booth, B. P., French, T. D., and Déry, S. J.: Determining contemporary and historical sediment sources in a large drainage basin impacted by cumulative effects: the regulated Nechako River, British Columbia, Canada, Journal of Soils and Sediments, 19, 3357–3373, https://doi.org/http://dx.doi.org/10.1007/s11368-019-02299-2, 2019.

Gellis, A. and Gorman Sanisaca, L.: Sediment Fingerprinting to Delineate Sources of Sediment in the Agricultural and Forested Smith Creek Watershed, Virginia, USA, JAWRA Journal of the American Water Resources Association, 54, 1197–1221, https://doi.org/10.1111/1752-1688.12680, 2018.

Gellis, A. C. and Noe, G. B.: Sediment source analysis in the Linganore Creek watershed, Maryland, USA, using the sediment fingerprinting approach: 2008 to 2010, Journal of Soils and Sediments, 13, https://doi.org/10.1007/s11368-013-0771-6, 2013.

Gellis, A. C. and Walling, D. E.: Sediment Source Fingerprinting (Tracing) and Sediment Budgets as Tools in Targeting River and Watershed Restoration Programs, in: Geophysical Monograph Series, edited by Simon, A., Bennett, S. J., and Castro, J. M., pp. 263–291, American Geophysical Union, Washington, D. C., https://doi.org/10.1029/2010GM000960, 2013.

Gibbs, M. M.: Identifying Source Soils in Contemporary Estuarine Sediments: A New Compound-Specific Isotope Method, Estuaries and Coasts, 31, https://doi.org/10.1007/s12237-007-9012-9, 2008.

Haddadchi, A., Olley, J., and Laceby, P.: Accuracy of mixing models in predicting sediment source contributions, Science of The Total Environment, 497-498, 139–152, https://doi.org/10.1016/j.scitotenv.2014.07.105, 2014.

Hollander, M.: DA Wolfe. 1973. Nonparametric statistical methods, John Wiley and Sons Perry, P. and S. Wolff, 1074, 156–158, 1973.

Horowitz, A. J.: A primer on sediment-trace element chemistry, vol. 2, Lewis Publishers Chelsea, https://doi.org/http://dx.doi.org/10.3133/ofr9176, 1991.

Huangfu, Y., Essington, M. E., Hawkins, S. A., Walker, F. R., Schwartz, J. S., and Layton, A. C.: Testing the sediment fingerprinting technique using the SIAR model with artificial sediment mixtures, Journal of Soils and Sediments, 20, 1771–1781, https://doi.org/10.1007/s11368-019-02545-7, 2020.

Huon, S., Hayashi, S., Laceby, J. P., Tsuji, H., Onda, Y., and Evrard, O.: Source dynamics of radiocesium-contaminated particulate matter deposited in an agricultural water reservoir after the Fukushima nuclear accident, Science of the Total Environment, 612, 1079–1090, https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.07.205, 2018.

Issaka, S. and Ashraf, M. A.: Impact of soil erosion and degradation on water quality: a review, Geology, Ecology, and Landscapes, 1, 1–11, https://doi.org/10.1080/24749508.2017.1301053, 2017.

JAXA: High-Resolution Land Use and Land Cover Map of Japan [2006-2011] (ver. 16.09 ; 10-m resoution ; 12 categories), https://www.eorc.jaxa.jp/ALOS/en/dataset/lulc/lulc_jpn_e.htm, 2016.

JAXA: Resolution Land Use and Land Cover Map of Japan [2014-2016] (ver. 18.03 ; 30-m resoution ; 12 categories), https://www.eorc.jaxa.jp/ALOS/en/dataset/lulc/lulc_v1803_e.htm, 2018.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

JAXA: High-Resolution Land Use and Land Cover Map of Japan [2018-2020] (ver. 21.11 ; 10-m resoution ; 12 categories), https://www.eorc.jaxa.jp/ALOS/en/dataset/lulc/lulc_v2111_e.htm, 2021.

Jordan, A., Krüger, F., and Lerch, S.: Evaluating probabilistic forecasts with scoringRules, arXiv preprint arXiv:1709.04743, https://doi.org/https://doi.org/10.48550/arXiv.1709.04743, 2017.

Jordan, A., Krüger, F., and Lerch, S.: Evaluating Probabilistic Forecasts with scoringRules, Journal of Statistical Software, 90, 1–37, https://doi.org/10.18637/jss.v090.i12, r package version 1.0.2, 2019.

Kemp, P., Sear, D., Collins, A., Naden, P., and Jones, I.: The impacts of fine sediment on riverine fish, Hydrological Processes, 25, 1800–1821, https://doi.org/10.1002/hyp.7940, 2011.

Koiter, A., Owens, P., Petticrew, E., and Lobb, D.: The Behavioural Characteristics of Sediment Properties and Their Implications for Sediment Fingerprinting as an Approach for Identifying Sediment Sources in River Basins, Earth-Science Reviews, 125, 24–42, https://doi.org/10.1016/j.earscirev.2013.05.009, 2013.

Koiter, A. J., Owens, P. N., Petticrew, E. L., and Lobb, D. A.: Assessment of particle size and organic matter correction factors in sediment source fingerprinting investigations: An example of two contrasting watersheds in Canada, Geoderma, 325, 195–207, https://doi.org/10.1016/j.geoderma.2018.02.044, 2018.

Laceby, J., Batista, P., Taube, N., Kruk, M., Chung, C., Evrard, O., Orwin, J., and Kerr, J.: Tracing total and dissolved material in a western Canadian basin using quality control samples to guide the selection of fingerprinting parameters for modelling, CATENA, 200, 105 095, https://doi.org/10.1016/j.catena.2020.105095, 2021a.

Laceby, J. P. and Olley, J.: An examination of geochemical modelling approaches to tracing sediment sources incorporating distribution mixing and elemental correlations, Hydrological processes, 29, 1669–1685, https://doi.org/https://doi.org/10.1002/hyp.10287, 2015.

Laceby, J. P., McMahon, J., Evrard, O., and Olley, J.: A comparison of geological and statistical approaches to element selection for sediment fingerprinting, Journal of Soils and Sediments, 15, 2117–2131, 2015.

Laceby, J. P., Chartin, C., Evrard, O., Onda, Y., Garcia-Sanchez, L., and Cerdan, O.: Rainfall erosivity in catchments contaminated with fallout from the Fukushima Daiichi nuclear power plant accident, Hydrology and Earth System Sciences, 20, 2467–2482, https://doi.org/https://doi.org/10.5194/hess-20-2467-2016, 2016a.

Laceby, J. P., Huon, S., Onda, Y., Vaury, V., and Evrard, O.: Do forests represent a long-term source of contaminated particulate matter in the Fukushima Prefecture?, Journal of Environmental Management, 183, 742–753, https://doi.org/https://doi.org/10.1016/j.jenvman.2016.09.020, 2016b.

Laceby, J. P., Evrard, O., Smith, H. G., Blake, W. H., Olley, J. M., Minella, J. P., and Owens, P. N.: The Challenges and Opportunities of Addressing Particle Size Effects in Sediment Source Fingerprinting: A Review, Earth-Science Reviews, 169, 85–103, https://doi.org/10.1016/j.earscirev.2017.04.009, 2017.

Laceby, J. P., Batista, P. V. G., Taube, N., Kruk, M. K., Chung, C., Evrard, O., Orwin, J. F., and Kerr, J. G.: Tracing total and dissolved material in a western Canadian basin using quality control samples to guide the selection of fingerprinting parameters for modelling, Catena, 200, 105 095, https://doi.org/https://doi.org/10.1016/j.catena.2020.105095, 2021b.

Laio, F. and Tamea, S.: Verification tools for probabilistic forecasts of continuous hydrological variables, Hydrology and Earth System Sciences, 11, 1267–1277, https://doi.org/http://dx.doi.org/10.5194/hess-11-1267-2007, 2007.

Lal, R.: Soil Erosion Impact on Agronomic Productivity and Environment Quality, Critical Reviews in Plant Sciences, 17, 319–464, https://doi.org/10.1080/07352689891304249, 1998.

Lal, R.: Soil degradation by erosion, Land Degradation & Development, 12, 519–539, https://doi.org/10.1002/ldr.472, 2001.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Lal, R.: Accelerated soil erosion as a source of atmospheric CO_2, Soil and Tillage Research, 188, 35–40, https://doi.org/https://doi.org/10.1016/j.still.2018.02.001, 2019.

Lamb, A. L., Wilson, G. P., and Leng, M. J.: A review of coastal palaeoclimate and relative sea-level reconstructions using δ13C and C/N ratios in organic material, Earth-Science Reviews, 75, 29–57, https://doi.org/https://doi.org/10.1016/j.earscirev.2005.10.003, 2006.

Latorre, B., Lizaga, I., Gaspar, L., and Navas, A.: A novel method for analysing consistency and unravelling multiple solutions in sediment fingerprinting, Science of The Total Environment, 789, 147 804, https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.147804, 2021.

Li, Z. and Fang, H.: Impacts of climate change on water erosion: A review, Earth-Science Reviews, 163, 94–117, https://doi.org/https://doi.org/10.1016/j.earscirev.2016.10.004, 2016.

Lizaga, I., Gaspar, L., Blake, W. H., Latorre, B., and Navas, A.: Fingerprinting changes of source apportionments from mixed land uses in stream sediments before and after an exceptional rainstorm event, Geomorphology, 341, 216–229, https://doi.org/10.1016/j.geomorph.2019.05.015, 2019.

Lizaga, I., Latorre, B., Gaspar, L., and Navas, A.: Consensus Ranking as a Method to Identify Non-Conservative and Dissenting Tracers in Fingerprinting Studies, Science of The Total Environment, 720, 137 537, https://doi.org/10.1016/j.scitotenv.2020.137537, 2020a.

Lizaga, I., Latorre, B., Gaspar, L., and Navas, A.: FingerPro: An R Package for Tracking the Provenance of Sediment, Water Resources Management, 34, 3879–3894, https://doi.org/10.1007/s11269-020-02650-0, 2020b.

Loughran, R. J., Campbell, B. L., and Walling, D. E.: Soil erosion and sedimentation indicated by caesium 137: Jackmoor Brook catchment, Devon, England, Catena, 14, 201–212, https://doi.org/https://doi.org/10.1016/S0341-8162(87)80018-8, 1987.

Martínez-Carreras, N., Udelhoven, T., Krein, A., Gallart, F., Iffly, J. F., Ziebel, J., Hoffmann, L., Pfister, L., and Walling, D. E.: The use of sediment colour measured by diffuse reflectance spectrometry to determine sediment sources: application to the Attert River catchment (Luxembourg), Journal of Hydrology, 382, 49–63, https://doi.org/http://dx.doi.org/10.1007/s11368-009-0162-1, 2010.

Martínez-Carreras, N., Gallart, F., Iffly, J., Pfister, L., Walling, D., Krein, A., et al.: Uncertainty assessment in suspended sediment fingerprinting based on tracer mixing models: a case study from Luxembourg, IAHS publication, 325, 94, 2008.

Martínez-Carreras, N., Krein, A., Udelhoven, T., Gallart, F., Iffly, J. F., Hoffmann, L., Pfister, L., and Walling, D. E.: A rapid spectral-reflectance-based fingerprinting approach for documenting suspended sediment sources during storm runoff events, Journal of Soils and Sediments, 10, https://doi.org/10.1007/s11368-009-0162-1, 2010.

Matheson, J. E. and Winkler, R. L.: Scoring rules for continuous probability distributions, Management science, 22, 1087–1096, https://doi.org/https://doi.org/10.1287/mnsc.22.10.1087, 1976.

Meybeck, M. and Helmer, R.: The quality of rivers: from pristine stage to global pollution, Palaeogeography, Palaeoclimatology, Palaeoecology, 75, 283–309, https://doi.org/https://doi.org/10.1016/0031-0182(89)90191-0, 1989.

Minella, J. P., Walling, D. E., and Merten, G. H.: Combining sediment source tracing techniques with traditional monitoring to assess the impact of improved land management on catchment sediment yields, Journal of Hydrology, 348, 546–563, https://doi.org/10.1016/j.jhydrol.2007.10.026, 2008.

Mingus, K. A., Liang, X., Massoudieh, A., and Jaisi, D. P.: Stable Isotopes and Bayesian Modeling Methods of Tracking Sources and Differentiating Bioavailable and Recalcitrant Phosphorus Pools in Suspended Particulate Matter, Environmental Science & Technology, 53, 69–76, https://doi.org/10.1021/acs.est.8b05057, 2019.

Montgomery, D. R.: Soil erosion and agricultural sustainability, Proceedings of the National Academy of Sciences, 104, 13 268–13 272, https://doi.org/https://doi.org/10.1073/pnas.0611508104, 2007.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Mukundan, R., Radcliffe, D. E., Ritchie, J. C., Risse, L. M., and McKinley, R. A.: Sediment Fingerprinting to Determine the Source of Suspended Sediment in a Southern Piedmont Stream, Journal of Environmental Quality, 39, 1328–1337, https://doi.org/10.2134/jeq2009.0405, 2010.

Mukundan, R., Walling, D. E., Gellis, A. C., Slattery, M. C., and Radcliffe, D. E.: Sediment Source Fingerprinting: Transforming From a Research Tool to a Management Tool, JAWRA Journal of the American Water Resources Association, 48, 1241–1257, https://doi.org/10.1111/j.1752-1688.2012.00685.x, 2012.

NARO: Comprehensive soil map of agricultural land at 1:50,000 scale (Shapefile by Prefecture and National), https://soil-inventory.rad.naro.go.jp/download5.html, 2011.

Nosrati, K., Govers, G., Semmens, B. X., and Ward, E. J.: A mixing model to incorporate uncertainty in sediment fingerprinting, Geoderma, 217-218, 173–180, https://doi.org/10.1016/j.geoderma.2013.12.002, 2014.

Nosrati, K., Mohammadi-Raigani, Z., Haddadchi, A., and Collins, A. L.: Elucidating intra-storm variations in suspended sediment sources using a Bayesian fingerprinting approach, Journal of Hydrology, 596, 126 115, https://doi.org/10.1016/j.jhydrol.2021.126115, 2021.

Obara, H., Ohkura, T., Takata, Y., Kohyama, K., Maejima, Y., Hamazaki, T., et al.: Comprehensive soil classification system of Japan first approximation., Nogyo Kankyo Gijutsu Kenkyusho Hokoku= Bulletin of National Institute for Agro-Environmental Sciences, pp. 1–73, 2011.

Obara, H., Maejima, Y., Kohyama, K., Ohkura, T., and Takata, Y.: Outline of the comprehensive soil classification system of Japan–first approximation, Japan Agricultural Research Quarterly: JARQ, 49, 217–226, https://doi.org/https://doi.org/10.6090/jarq.49.217, 2015.

OCC: Soil is a non-renewable resource, FAO, p. 4, https://www.fao.org/documents/card/en/c/ec28fc04-3d38-4e35-8d9b-e4427e20a4f7/, 2015.

Olson, K. R., Al-Kaisi, M., Lal, R., and Cihacek, L.: Impact of soil erosion on soil organic carbon stocks, Journal of Soil and water Conservation, 71, 61A–67A, https://doi.org/http://dx.doi.org/10.2489/jswc.71.3.61A, 2016.

Owens, P. N., Blake, W. H., Gaspar, L., Gateuille, D., Koiter, A., Lobb, D. A., Petticrew, E. L., Reiffarth, D., Smith, H. G., and Woodward, J.: Fingerprinting and tracing the sources of soils and sediments: Earth and ocean science, geoarchaeological, forensic, and human health applications, Earth-Science Reviews, 162, 1–23, https://doi.org/https://doi.org/10.1016/j.earscirev.2016.08.012, 2016.

Palazón, L., Latorre, B., Gaspar, L., Blake, W. H., Smith, H. G., and Navas, A.: Comparing catchment sediment fingerprinting procedures using an auto-evaluation approach with virtual sample mixtures, Science of The Total Environment, 532, https://doi.org/10.1016/j.scitotenv.2015.05.003, 2015.

Peart, M. R. and Walling, D. E.: Fingerprinting sediment source: the example of a drainage basin in Devon, UK, Drainage basin sediment delivery, 159, 41–55, https://doi.org/http://hdl.handle.net/10722/157753, 1986.

Phillips, I. and Greenway, M.: Changes in water-soluble and exchangeable ions, cation exchange capacity, and phosphorus$_{max}$ in soils under alternating waterlogged and drying conditions, Communications in soil science and plant analysis, 29, 51–65, https://doi.org/https://doi.org/10.1080/00103629809369928, 1998.

Pimentel, D.: Soil erosion: a food and environmental threat, Environment, development and sustainability, 8, 119–137, https://doi.org/http://dx.doi.org/10.1007/s10668-005-1262-8, 2006.

Pimstein, A., Notesco, G., and Ben-Dor, E.: Performance of Three Identical Spectrometers in Retrieving Soil Reflectance under Laboratory Conditions, Soil Science Society of America Journal, 75, 746–759, https://doi.org/10.2136/sssaj2010.0174, 2011.

Poesen, J.: Soil erosion in the Anthropocene: Research needs, Earth Surface Processes and Landforms, 43, 64–84, https://doi.org/https://doi.org/10.1002/esp.4250, 2018.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

QGIS Development Team: QGIS Geographic Information System, Open Source Geospatial Foundation, https://www.qgis.org/fr/site/, version 3.26.0, 2022.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, version 4.1.2, 2021.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, version 4.1.2, 2022.

RStudio Team: RStudio: Integrated Development Environment for R, RStudio, PBC, Boston, MA, http://www.rstudio.com/, version 2022.7.1.554, 2022.

Russi, T., Packard, A., Feeley, R., and Frenklach, M.: Sensitivity Analysis of Uncertainty in Model Prediction, The Journal of Physical Chemistry A, 112, 2579–2588, https://doi.org/10.1021/jp076861c, 2008.

Sellier, V.: Développement de méthodes de traçage sédimentaire pour quantifier l'impact des mines de nickel sur l'hyper-sédimentation des rivières et l'envasement des lagons de Nouvelle-Calédonie, PhD Thesis, 2020.

Sellier, V., Navratil, O., Laceby, J. P., Legout, C., Foucher, A., Allenbach, M., Lefèvre, I., and Evrard, O.: Combining colour parameters and geochemical tracers to improve sediment source discrimination in a mining catchment (New Caledonia, South Pacific Islands), SOIL, 7, 743–766, https://doi.org/10.5194/soil-7-743-2021, 2021.

Sherriff, S. C., Franks, S. W., Rowan, J. S., Fenton, O., and Ó'hUallacháin, D.: Uncertainty-based assessment of tracer selection, tracer non-conservativeness and multiple solutions in sediment fingerprinting using synthetic and field data, Journal of Soils and Sediments, 15, 2101–2116, https://doi.org/10.1007/s11368-015-1123-5, 2015.

Small, I. F., Rowan, J. S., Franks, S. W., Wyatt, A., and Duck, R. W.: Bayesian sediment fingerprinting provides a robust tool for environmental forensic geoscience applications, Geological Society, London, Special Publications, 232, 207–213, https://doi.org/https://doi.org/10.1144/GSL.SP.2004.232.01.19, 2004.

Smith, H. G. and Blake, W. H.: Sediment fingerprinting in agricultural catchments: A critical re-examination of source discrimination and data corrections, Geomorphology, 204, https://doi.org/10.1016/j.geomorph.2013.08.003, 2014.

Smith, H. G., Karam, D. S., and Lennard, A. T.: Evaluating tracer selection for catchment sediment fingerprinting, Journal of Soils and Sediments, 18, 3005–3019, https://doi.org/http://dx.doi.org/10.1007/s11368-018-1990-7, 2018.

Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., Macdonald, L. M., and McLaughlin, M. J.: The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties, Applied spectroscopy reviews, 49, 139–186, https://doi.org/https://doi.org/10.1080/05704928.2013.811081, 2014.

Stock, B., Semmens, B., Ward, E., Parnell, A., Jackson, A., and Phillips, D.: MixSIAR: Bayesian Mixing Models in R, https://doi.org/10.5281/zenodo.1209993, r package version 3.1.12, 2020.

Stock, B., Semmens, B., Ward, E., Parnell, A., Jackson, A., and Phillips, D.: JAGS: Bayesian Mixing Models in R, https://mcmc-jags.sourceforge.io/, r package version 4.3.1, 2022.

Summers, D., Lewis, M., Ostendorf, B., and Chittleborough, D.: Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties, Ecological Indicators, 11, https://doi.org/10.1016/j.ecolind.2009.05.001, 2011.

Tiecher, T., Caner, L., Minella, J. P. G., and dos Santos, D. R.: Combining visible-based-color parameters and geochemical tracers to improve sediment source discrimination and apportionment, Science of the Total Environment, 527, 135–149, https://doi.org/https://doi.org/10.1016/j.scitotenv.2015.04.103, 2015.

https://doi.org/10.5194/egusphere-2023-1970
Preprint. Discussion started: 8 September 2023

Tiecher, T., Moura-Bueno, J. M., Caner, L., Minella, J. P., Evrard, O., Ramon, R., Naibo, G., Barros, C. A., Silva, Y. J., Amorim, F. F., et al.: Improving the quantification of sediment source contributions using different mathematical models and spectral preprocessing techniques for individual or combined spectra of ultraviolet–visible, near-and middle-infrared spectroscopy, Geoderma, 384, 114 815, https://doi.org/https://doi.org/10.1016/j.geoderma.2020.114815, 2021.

Vale, S., Swales, A., Smith, H. G., Olsen, G., and Woodward, B.: Impacts of tracer type, tracer selection, and source dominance on source apportionment with sediment fingerprinting, Science of The Total Environment, 831, 154 832, https://doi.org/10.1016/j.scitotenv.2022.154832, 2022.

Viparelli, E., Lauer, J. W., Belmont, P., and Parker, G.: A numerical model to develop long-term sediment budgets using isotopic sediment fingerprints, Computers & geosciences, 53, 114–122, https://doi.org/https://doi.org/10.1016/j.cageo.2011.10.003, 2013.

Viscarra Rossel, R., Minasny, B., Roudier, P., and McBratney, A.: Colour space models for soil science, Geoderma, 133, 320–337, https://doi.org/10.1016/j.geoderma.2005.07.017, 2006.

Walden, J., Slattery, M., and Burt, T.: Use of mineral magnetic measurements to fingerprint suspended sediment sources: approaches and techniques for data analysis, Journal of Hydrology, 202, 353–372, https://doi.org/10.1016/S0022-1694(97)00078-4, 1997.

Wall, G. J. and Wilding, L. P.: Mineralogy and Related Parameters of Fluvial Suspended Sediments in Northwestern Ohio, Journal of Environmental Quality, 5, 168–173, https://doi.org/10.2134/jeq1976.00472425000500020012x, 1976.

Walling, D. and Woodward, J.: Use of radiometric fingerprints to derive information on suspended sediment sources, in: Erosion and sediment monitoring programmes in river basins. Proc. international symposium, Oslo, 1992l, pp. 153–164, International Association of Hydrological Sciences, https://www.researchgate.net/publication/239921785_Use_of_radiometric_fingerprints_to_derive_information_on_suspended_sediment_curves, 1992.

Walling, D. E., Owens, P. N., Waterfall, B. D., Leeks, G. J., and Wass, P. D.: The particle size characteristics of fluvial suspended sediment in the Humber and Tweed catchments, UK, Science of the Total Environment, 251, 205–222, https://doi.org/https://doi.org/10.1016/S0048-9697(00)00384-3, 2000.

Wei, P., Lu, Z., and Song, J.: Variable importance analysis: A comprehensive review, Reliability Engineering & System Safety, 142, 399–432, https://doi.org/10.1016/j.ress.2015.05.018, 2015.

Weihs, C., Ligges, U., Luebke, K., and Raabe, N.: klaR Analyzing German Business Cycles, in: Data Analysis and Decision Support, edited by Baier, D., Decker, R., and Schmidt-Thieme, L., pp. 335–343, Springer-Verlag, Berlin, r package version 1.7.1, 2005.

Whitaker, A. C., Chapasa, S. N., Sagras, C., Theogene, U., Veremu, R., and Sugiyama, H.: Estimation of baseflow recession constant and regression of low flow indices in eastern Japan, Hydrological Sciences Journal, 67, 191–204, 2022.

Wilkinson, S. N., Hancock, G. J., Bartley, R., Hawdon, A. A., and Keen, R. J.: Using sediment tracing to assess processes and spatial patterns of erosion in grazed rangelands, Burdekin River basin, Australia, Agriculture, Ecosystems & Environment, 180, 90–102, 2013.

Xu, Z., Belmont, P., Brahney, J., and Gellis, A. C.: Sediment source fingerprinting as an aid to large-scale landscape conservation and restoration: A review for the Mississippi River Basin, Journal of Environmental Management, 324, 116 260, https://doi.org/10.1016/j.jenvman.2022.116260, 2022.