Science of the Total Environment 532 (2015) 456–466

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Comparing catchment sediment fingerprinting procedures using an auto-evaluation approach with virtual sample mixtures

Leticia Palazón [a,*], Borja Latorre [a], Leticia Gaspar [c], William H. Blake [b], Hugh G. Smith [d], Ana Navas [a]

[a] Department of Soil and Water, Estación Experimental de Aula Dei (EEAD-CSIC), Avda. Montañana 1005, Zaragoza, 50059, Spain
[b] School of Geography, Earth and Environmental Sciences, Plymouth University, Portland Square, Drake Circus, Plymouth, Devon, PL4 8AA, United Kingdom
[c] Environmental Science Program, University of Northern British Columbia, 3333 University Way, Prince George, BC V2N 4Z9, Canada
[d] School of Environmental Sciences, University of Liverpool, Liverpool L697ZT, United Kingdom

HIGHLIGHTS

- Virtual sample mixtures were generated from possible sediment sources.
- 24 fingerprinting procedures were tested.
- Goodness of fit should not be used as an accuracy index of contribution estimates.
- More tracers in a composite fingerprint would improve source apportionment results.
- Different contributions can be obtained with different fingerprinting procedures.

ARTICLE INFO

Article history:
Received 24 November 2014
Received in revised form 1 May 2015
Accepted 1 May 2015
Available online 20 June 2015

Editor: D. Barcelo

Keywords:
Sediment fingerprinting
Mixing model
Sediment source ascription
Sediment contribution
River catchments

ABSTRACT

Information on sediment sources in river catchments is required for effective sediment control strategies, to understand sediment, nutrient and pollutant transport, and for developing soil erosion management plans. Sediment fingerprinting procedures are employed to quantify sediment source contributions and have become a widely used tool. As fingerprinting procedures are naturally variable and locally dependant, there are different applications of the procedure. Here, the auto-evaluation of different fingerprinting procedures using virtual sample mixtures is proposed to support the selection of the fingerprinting procedure with the best capacity for source discrimination and apportionment. Surface samples from four land uses from a Central Spanish Pyrenean catchment were used i) as sources to generate the virtual sample mixtures and ii) to characterise the sources for the fingerprinting procedures. The auto-evaluation approach involved comparing fingerprinting procedures based on four optimum composite fingerprints selected by three statistical tests, three source characterisations (mean, median and corrected mean) and two types of objective functions for the mixing model. A total of 24 fingerprinting procedures were assessed by this new approach which were solved by Monte Carlo simulations and compared using the root mean squared error (RMSE) between known and assessed source ascriptions for the virtual sample mixtures. It was found that the source ascriptions with the highest accuracy were achieved using the corrected mean source characterisations for the composite fingerprints selected by the Kruskal Wallis *H*-test and principal components analysis. Based on the RMSE results, high goodness of fit (GOF) values were not always indicative of accurate source apportionment results, and care should be taken when using GOF to assess mixing model performance. The proposed approach to test different fingerprinting procedures using virtual sample mixtures provides an enhanced basis for selecting procedures that can deliver optimum source discrimination and apportionment.

1. Introduction

Research on source fingerprinting procedures and their development to provide information on the sources of sediment transported through a river catchment can be traced back to the 1970s including works of Klages and Hsieh (1975), Wall and Wilding (1976) and Walling et al. (1979). Since these early works, sediment source fingerprinting applications have expanded greatly. Walling (2013) identified a key driver behind the expansion of such work as the need to support the development of sediment management strategies aimed at dealing with environmental problems associated with fine sediment. This expansion in sediment fingerprinting procedure led to the use of variable sediment fingerprinting applications tailored to the wide range of

Abbreviations: GOF, goodness of fit; RMSE, root mean squared error; KW, Kruskal–Wallis *H*-test; DFA, discriminant function analysis; PCA, principal components analysis.

* Corresponding author.
E-mail address: lpalazon@eead.csic.es (L. Palazón).

http://dx.doi.org/10.1016/j.scitotenv.2015.05.003

potential controls on sediment properties and the contributions from catchment sediment sources. Sediment fingerprinting procedures offer potential to quantify the contribution of different catchment sediment sources, evaluate erosion dynamics and serve as a basis to develop management plans to tackle erosion and sediment related problems, especially in catchments with land use conflicts (Pacheco et al., 2014; Valle Junior et al., 2014).

Based on differences in source material properties, fine sediment fingerprinting allows the discrimination and apportionment of sediment derived from sampled catchment sources (Walling et al., 1999). The use of statistical tests to confirm the ability of the properties to discriminate between the sources and to select the best subset of properties for the composite fingerprint in most early fingerprinting studies were unnecessary as they were based in a limited number of sources (e.g., two) and tracer properties (perhaps only one). Along with the development of the fingerprinting procedure, the number of potential sources and fingerprint properties increased and, therefore, the need to use statistical tests to select the optimum composite fingerprints became more important and therefore was increasingly recognized. As a minimum, $n - 1$ properties are necessary to discriminate rigorously between n sources. Additional properties are frequently necessary to increase the reliability of the results (Walling, 2013). These tracer properties may include geochemical, radionuclide, mineral magnetic, organic constituent, stable isotope and colour properties (Foster and Lees, 2000). Therefore, the sediment fingerprinting procedure typically first identifies a subset of tracer properties that discriminate the sampled sources by different statistical tests (Collins and Walling, 2002) and then estimates the proportional contributions from each source using mixing models to solve the set of linear equations characterised by the selected tracer properties (e.g., Yu and Oldfield, 1989; Motha et al., 2003; Martínez-Carreras et al., 2010; Blake et al., 2012; Owens et al., 2012; Schuller et al., 2013; Smith and Blake, 2014). Source apportionments are obtained by the solution of a set of linear equations characterised by an objective function, which represents the relation between a tracer property value in sediment with the sum of multiplications between that tracer value and the unknown apportionment for each source by optimization approaches. Several variants of the objective function have been used by different authors by incorporating correction factors for differences in particle size and organic matter content between target and source material samples (Collins et al., 1997) and the use of weightings and elemental correlations for the individual tracer properties, in order to vary the emphasis placed on individual properties when fitting the model (e.g., Collins et al., 2010, 2012; Laceby and Olley, 2014). Although most fingerprinting studies employed local optimization routines to obtain the source contributions, these routines can fail to find the best optimum solution (Collins et al., 2012). Genetic algorithm optimization and the use of stratified random sampling of the property probability distributions using Latin Hypercube Sampling have been proposed to overcome this problem (e.g., Collins et al., 2012; Haddadchi et al., 2013). Other tools such as Bayesian approaches in mixing model applications have also been successfully applied in fingerprinting procedures (e.g., Fox and Papanicolaou, 2008; Massoudieh et al., 2012; D'Haen et al., 2013).

There is a range of different applications of the sediment fingerprinting procedure in the literature and, in general, the greatest methodological differences are related to i) the statistical analysis used to identify the subset of the tracer properties which discriminate between sources; ii) the way in which the sources were characterised for the mixing model (i.e., mean, median or corrected mean); iii) the use of correction factors (including weighting and elemental correlations) in the objective function; iv) the type of the objective function and v) the optimisation procedure used to solve the mixing model. These differences between applications were in many cases due to the specific characteristics of the study areas and, therefore, the selection of the most effective fingerprinting procedure for each specific application can become time-consuming and complex.

The accuracy and sensitivity of the tracer selection and source unmixing procedures associated with sediment fingerprinting have received limited attention. Haddadchi et al. (2013) compared mixing models applying local and global optimization methods to datasets from two different catchments and indicated that the mixing model outputs could change remarkably depending on which mixing model was used. More recently, Haddadchi et al. (2014) compared the accuracy of four defined mixing models to solve artificial mixture samples from three well-differentiated sources concluding that there is a need to test mixing models using known source and mixture samples prior to applying them to field samples. Laceby and Olley (2014) compared different mixing models used in the literature to analyse artificial mixture samples based on catchment sources and concluded that the most accurate procedures incorporated correlations between elements and did not use tracer discriminatory weightings. These few studies highlight the methodological uncertainty hampering the wider adoption of the fingerprinting approach for identifying sediment sources. There remains a need for further methodological guidance to aid the assessment of accuracy and to support pre-selection of the most effective fingerprinting procedures for catchment applications.

Whereas the accuracy of the fingerprinting procedures has started to be evaluated with well-differentiated sources (Haddadchi et al., 2014), this study aims to evaluate the accuracy of a set of fingerprinting procedures for a river catchment in which sources are less well differentiated. The selected catchment is representative of the Mediterranean environment that was subject to intense land use changes that drive sediment production and where sediment sources based on land use might not be clearly discriminated. As an approach for pre-selecting the most effective fingerprinting procedure, we propose to test the discriminatory accuracy of different fingerprinting procedures by generating virtual sample mixtures using known and natural source samples. These virtual sample mixtures were used to assess the capacity of various fingerprinting procedures to reproduce the known source apportionments. The auto-evaluation of the procedure could guide the fingerprinting procedure design and be used to assess the robustness of the results.

2. Material and methods

2.1. Study area

The samples used to characterise potential sources and create the virtual sample mixtures were collected in the Isábena River catchment (445 km^2) of the Central Spanish Pyrenees (Fig. 1). Climatically, the catchment falls in the Mediterranean domain. Mean annual precipitation in the catchment is around 767 mm and ranges from 450 mm at the outlet to 1600 mm at the headwater (Verdú et al., 2006). The mean annual temperature ranges from 12.5 °C at the outlet to 10 °C in the headwater. The headwater of the catchment is partially karstified with predominance of Cretaceous limestones. In the intermediate part of the catchment the presence of Eocene marls comprises depressions in which badlands are developed. In the southern lowland area, Tertiary sedimentary rocks (clays, sandstones and conglomerates) are predominant. The climatic and topographic characteristics of the catchment influenced the distribution of land uses in the Isábena catchment. Therefore, the agricultural lands predominate in the lowland areas, whereas forests and interspersed grasslands and scrubland dominate the highlands (Fig. 1). Forests and grassland are the main land uses occupying more than 50% of the catchment, followed by cultivated land that occupies less than 20% and scrublands which cover 10% of the catchment surface area. Important changes in land use occurred during the last 60 years in the Spanish Pyrenean region, resulting in substantial land abandonment that has affected most parts of the agricultural areas triggering the subsequent process of natural reforestation (Navas et al., 2008).