This work has been submitted to NECTAR, the Northampton Electronic Collection of Theses and Research.

**Article**

**Title:** 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

**Creators:** Collins, A. L., Pulley, S., Foster, I. D. L., Gellis, A., Porto, P. and Horowitz, A. J.

**DOI:** 10.1016/j.jenvman.2016.09.075


It is advisable to refer to the publisher's version if you intend to cite from this work.

**Version:** Published version

**Official URL:** http://dx.doi.org/10.1016/j.jenvman.2016.09.075

**Note:**

![Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/)

This work is licensed under a Creative Commons Attribution 3.0 Unported License.

http://nectar.northampton.ac.uk/8839/
Sediment article

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

A.L. Collins a,*, S. Pulley b, I.D.L. Foster b, c, A. Gellis d, P. Porto e, A.J. Horowitz f,1

a Sustainable Soils and Grassland Systems Department, Rothamsted Research, Okehampton, EX20 2SB, UK
b Geography Department, Rhodes University, Grahamstown, 6140, South Africa
c School of Science and Technology, University of Northampton, Northampton, NN2 6JD, UK
d U.S. Geological Survey, Baltimore, MD, 21223, United States
e Department of Agraria, University Mediterranean of Reggio Calabria, Italy
f U.S. Geological Survey, Atlanta, GA, 30093, United States

ARTICLE INFO

Article history:
Received 12 July 2016
Received in revised form 20 September 2016
Accepted 22 September 2016
Available online xxx

Keywords:
Suspended sediment sources
Source fingerprinting
Source tracing
Mixing model
Uncertainty

ABSTRACT

The growing awareness of the environmental significance of fine-grained sediment fluxes through catchment systems continues to underscore the need for reliable information on the principal sources of this material. Source estimates are difficult to obtain using traditional monitoring techniques, but sediment source fingerprinting or tracing procedures, have emerged as a potentially valuable alternative. Despite the rapidly increasing numbers of studies reporting the use of sediment source fingerprinting, several key challenges and uncertainties continue to hamper consensus among the international scientific community on key components of the existing methodological procedures. Accordingly, this contribution reviews and presents recent developments for several key aspects of fingerprinting, namely: sediment source classification, catchment source and target sediment sampling, tracer selection, grain size issues, tracer conservatism, source apportionment modelling, and assessment of source predictions using artificial mixtures. Finally, a decision-tree representing the current state of knowledge is presented, to guide end-users in applying the fingerprinting approach.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction to sediment source fingerprinting

Reliable quantitative information on fine-grained sediment sources in river catchments is required to help target remedial actions for mitigating the impacts of excessive fine sediment loss on aquatic biology (Kemp et al., 2011; Jones et al., 2012). Such knowledge can also help reduce the contribution of high sediment loads to drinking water treatment costs (Lal and Stewart, 2013), the maintenance of water storage reservoirs (Verstraeten and Poesen, 2000), and navigation routes (Milliman and Meade, 1983). The role of fine sediment redistribution as a key vector for the transfer of nutrients and contaminants (Horowitz, 1985; Allan, 1986) across the land-to-water continuum has also been a key driver for the increased need for information on fine-grained sediment provenance. The use of tracers to infer fine-grained (typically <63 µm) sediment provenance qualitatively dates back to the 1970s (Klages and Hsieh, 1975; Wall and Wilding, 1976; Walling et al., 1979). As the popularity of such approaches began to increase, statistical methods were introduced to improve the robustness of source discrimination (Yu and Oldfield, 1989; Walling and Woodward, 1995; Collins et al., 1996, 1997a). In addition, mathematical un-mixing modelling was introduced for the quantitative apportionment of sediment provenance (Walling et al., 1993; Walling and Woodward, 1995; Collins et al., 1996, 1997a) because it soon became apparent that no single tracer could discriminate robustly between multiple potential sediment sources. This realization also resulted in the growing application of composite signatures combining tracers with differing environmental controls (Walling et al., 1993; Collins et al., 1997a; Devereux et al., 2010). Accordingly, a wide range of tracer properties have been tested and applied in the growing body of studies using the fingerprinting approach (Collins and Walling, 2004; Walling, 2005, 2013;
Haddadchi et al., 2013; Guzmán et al., 2013; Miller et al., 2015; Collins, 2015). The physical properties tested include colour ( Grimshaw and Lewin, 1980; Krein et al., 2003; Croft and Pye, 2004; Martinez-Carreras et al., 2010; Barthod et al., 2015) and grain size (Kurashige and Fusejima, 1997; Weltje and Prins, 2003, 2007; Weltje, 2012). Chemical properties include clay mineralogy (Eberl, 2004; Gingele and De Deckker, 2005), mineral-magnetism (Yu and Oldfield, 1993; Caithcheon, 1998; Maher et al., 2009; Dearing, 2000; Zhang et al., 2008; Hatfield and Maher, 2009), geochemistry (Collins and Walling, 2002), fallout radionucleides (Wallbrink and Murray, 1993; Krause et al., 2003; Wilkinson et al., 2013; Belmont et al., 2014; Evrard et al., 2016), cosogenic radionuclides (Perg et al., 2003), bulk stable isotopes and isotopic ratios (Yang et al., 2008; Fox and Papanicolaou, 2008), and biomarkers (Hancock and Revill, 2013; Alewell et al., 2016; Reiffarth et al., 2016). Biological properties include soil enzymes (Nosrati et al., 2011) and pollen (Brown, 1985). The underlying assumption in the application of these various tracer groups is that they provide a robust basis for discriminating potential sediment sources, although in many instances, source discrimination is tested by finding a statistical solution using either parametric (Collins et al., 2010a) or Bayesian (Stewart et al., 2014) approaches. Inherent in the sediment fingerprinting approach are the additional assumptions that the tracer properties are measurable, conservative (e.g. don’t change from source to sink or evolve in a predictable manner), and representative. These assumptions have been and continue to be scrutinized (e.g. Foster and Lees, 2000; Koiter et al., 2013) and represent an area of much needed further research.

Despite the recent growing application of sediment source tracing (Walling, 2013; Guzmán et al., 2013; Haddadchi et al., 2013; Miller et al., 2015; Walling and Collins, 2016; Walling and Foster, 2016), there remains a strategic need to continue refining and, perhaps more importantly, standardizing the procedures therein. In response to the many questions being raised in this growing field, a technical workshop was organised by the International Commission on Continental Erosion (ICCE) at the 26th International Union of Geodesy and Geophysics meeting in 2015 in Prague to review methodological aspects of sediment fingerprinting, thus resulting in this special section. The following sections provide a brief overview of some of the issues discussed at that workshop and which are explored in the papers in this volume. The sections herein explore progress and remaining issues related to a number of fundamental steps required for the successful application of sediment source tracing including: source classification and sampling, target sediment collection, tracer selection, grain size considerations, tracer conservatism, source discrimination, and apportionment modelling and evaluation. The paper culminates in the presentation of a new decision-tree designed to guide end-users through a series of critical decisions needed to apply the fingerprinting approach to apportion fine-grained sediment sources in river catchments. This decision-tree builds on earlier versions of methodological flow charts including those presented in Lees (1999), Foster and Lees (2000), Walling and Collins (2000), Collins and Walling (2004), and Walling et al. (2003a, 2006), and critically, captures both historical and recent research experiences and lessons.

2. Sediment source classification

A key consideration in the application of sediment fingerprinting relates to the classification of potential catchment sources. The fundamental distinction (Collins and Walling, 2004) made here concerns individual source types (e.g. surface or land use-based versus subsurface i.e., stream banks) and spatial (e.g. geological units or tributary sub-catchments) sediment sources. Some recent work has combined traditional sediment fingerprinting with particulate tracking techniques to increase the resolution of land use-based source types (Collins et al., 2010a, 2013a). In some cases, the classification of sediment source types has been founded on the primary processes of sediment generation including mass wasting, and sheet, rill and gully erosion (Wallbrink and Murray, 1993; Gellis et al., 2009; Gellis and Walling, 2011; Miller et al., 2015); whereas others have combined sediment source types and spatial units (e.g. Collins et al., 1997b; Walling et al., 1999; Juracek and Ziegler, 2009; Wilkinson et al., 2009).

Classification of sediment source groups is most commonly performed a priori to align source apportionment estimates with land use patterns and corresponding management goals (e.g. Peart and Walling, 1986; Walling and Woodward, 1995; Collins et al., 1997a; Owens et al., 1999; Porto et al., 2005; Collins et al., 2010a,b,c,d; Smith and Blake, 2014; Lamba et al., 2015; Foucher et al., 2015). Classification by land use has clear practical advantages, as sediment loss assigned to cultivated or pasture land within a catchment, for example, can be targeted by relevant management strategies based on on-farm interventions such as minimum tillage or compaction management. Tracer concentrations in a land use-based source group are likely to be controlled by numerous factors including soil and colluvium parent material, pedogenic processes, anthropogenic inputs, or prevailing soil moisture conditions (e.g. gleying). These factors have the potential to increase within-source variability, with clear implications for the uncertainty ranges associated with predicted source apportionment. The fundamental requirement to reduce intra-group and increase intergroup tracer variability (Small et al., 2002; Collins and Walling, 2002; Pulley et al., 2015a) is likely to be complicated or even confounded if small differences exist in tracer concentrations between different land use or surface and subsurface sources. Small differences in tracer concentrations can be caused by, for example, pedogenic processes or anthropogenic tracer inputs (e.g. atmospheric fallout of particulate lead from combustion). Based on historic land use and/or the chemical properties of the sources, different land uses can be combined, such as in the case of combining pasture with cropland to produce a general source group called agriculture (Gellis et al., 2015; Collins, 2015). Additional complexities must be considered where the fingerprinting approach is used to reconstruct sediment sources through time using sedimentary deposits. Here, some tracers cannot be assumed to have remained constant (behaved conservatively) through time due, for example, to atmospheric pollution elevating concentrations in catchment topsoil sources, as is likely to be the case for heavy metals and nutrients in artificial fertilisers (Foster and Charlesworth, 1996; Foster and Lees, 2000) or as the result of post-depositional dissolution and remobilization.

Geology has commonly been used as a basis of spatial source classification, both alone (e.g. D’Haen et al., 2012; Lacey and Olley, 2015), or paired with land use based classification (e.g. Collins et al., 1998; Owens et al., 1999). Classification by geology may provide less useful information for management purposes as the areas of a catchment underlain by a specific geology may be scattered irregularly around a catchment or a catchment may have homogenous geology under which circumstances an alternative source classification scheme would be warranted. In many catchments, however, geology may provide a convenient basis for classifying different regions such as steep hillslopes and uplands used for grazing as opposed to valley bottoms utilised for intensive cultivation and habitation (e.g. Collins et al., 1998; van der Waal et al., 2015). In the latter situation, whilst the use of stratigraphic units aids source discrimination, there is a clear need to identify the major sediment generation processes within those source groups to ensure that management interventions are targeted.
Objective sediment source grouping, as opposed to a priori
determination, especially in the context of minimising uncertainty
associated with quantitative source apportionment estimates, has
also been identified as an area of research worthy of more attention
(see Pulley et al., 2016 this issue). Pulley et al., 2016 (this issue)
expand on a method introduced by Walling et al. (1993) based on
pre-selected tracers from cluster analysis to classify sediment
source groups. Walling et al. (1993) found that land use was the
primary controlling factor on the resulting four to six source
groups. Walling and Woodward (1995) also used cluster analysis to
classify source groups according to pre-selected tracers. In contrast
to Walling et al. (1993), geology was identified as the major con-
trolling factor affecting source group classification, presumably
reflecting the larger and geologically more diverse drainage basin
used in the latter study. Using a similar cluster analysis method,
Pulley et al. (2016 this issue) used the tracer signatures of the
source samples in a PCA and cluster analysis to select the source
grouping which best fits the measured tracer signatures and this
was combined with the modification of the cluster groupings to suit
management goals. Wilkinson et al. (2015) modelled soil erosion and
collected source fingerprinting samples with priority given to
heavily eroding areas with the aim of reducing the potential im-
pacts of within-source group variability and making the source
samples more representative of heavily eroding sediment sources.
Similarly, van der Wal et al. (2015) retrieved source samples from
key erosional features, such as gullies, which were identified using
aerial photography. These approaches do, however, require a clear
understanding of the connectivity between eroding areas and the
river channel, which continues to be a fundamental premise for
targeting source sampling as eroded sediment will often not reach
the river channel in short time periods and contributing areas will
vary during runoff events of different magnitude and frequency
(Fryirs, 2013). Few source tracing studies publish maps of actual
source-to-river connectivity, but many (e.g. Gellis and Noe, 2013; Lacey et al., 2015) publish source sampling location maps (as
opposed to just geology and land use maps) which can be assumed
to be indicative of such connectivity given the underpinning
assumption in applying fingerprinting procedures that active
sources are sampled.

3. Catchment source sampling

Sampling protocols for catchment sources continue to be refined
with recent developments including the combination of indepen-
dent lines of evidence with more traditional strategies. Here, for
example, some researchers have combined process-based model
characterisation of catchments to help target source sampling
(Wilkinson et al., 2015; Theuring et al., 2015). Geomorphic assess-
ments (Wethered et al., 2015) and Google Earth (Boardman, 2016)
help can ensure that source sampling strategies are better
informed. Source material sampling is most commonly conducted
during a single intensive campaign. For many tracers this is an
adequate means of sampling; however, some organic tracers are
strongly influenced by seasonality. For example, nutrients from
decomposing post-reproductive salmon carcasses can be a signifi-
cant seasonal contribution to aquatic organic matter in some rivers
(Bilby et al., 1996; Ben-David et al., 1998). Sediment-associated
organic matter in streams is primarily composed of bacteria,
algae, soil-derived organic matter, macrophytes, leaf detritus, and
human septic waste and these sources can have seasonal patterns
of readily available material associated with, for example, vegeta-
tion dieback. In the absence of repeat source sampling campaigns
over time, sediment-associated organic matter sources can be
traced reliably, but for constrained time periods that do not include
all seasons (Collins et al., 2013b; 2014). Given these issues, source
sample scheduling must either take into account seasonal issues
associated with organic tracers, or alternatively, eliminate those
tracers that are subject to substantial intra-annual variability. A
similar issue is associated with agricultural crop rotations, that are
common in lowland intensive agricultural landscapes, and which
have been the feature of many sediment source tracing studies.
Crop rotation in particular degrades the strong discrimination be-
tween arable and grassland surface soils as potential sediment
sources provided by fallout radionuclides (e.g. Cs-137, unsupported
Pb-210, Be-7) by generating more overlap between the measured
concentrations to the depth of the plough layer. This complication
is less evident for unsupported Pb-210 or Be-7 since fallout is
ongoing, thereby providing a means of re-setting cultivation effects
even in the context of down-profile transfers over time. However,
re-setting of down-profile contrasts in fallout nuclide signatures,
for example, will depend on nuclide half life; Be-7 (half life 53 days)
will return to its pre-disturbance profile form much more quickly
than Pb-210 (half life ~ 22 years). More research is required to fully
understand the effects of land use change on down profile radio-
nuclide, physical, chemical and magnetic signatures.

When tracing the sources of historically deposited sediment it
also must be borne in mind that the concentrations of many tracers
in source groups have the propensity to change over long time
scales (decades and longer). Examples of tracers which may vary
in concentration in source groups over long timescales include
phosphorus, which is applied in artificial fertilisers, or lead and
other heavy metals where atmospheric releases, through produc-
tion, processing or consumption, have changed over time with
increasing industrialisation and/or regulation (Foster and
Charlesworth, 1996). Therefore, relevant timescales in the context
of the temporal stability (conservatism) of tracers used to distin-
guish catchment source groupings over long time spans must be
borne in mind (Foster and Lees, 2000).

Many sediment fingerprinting studies address sources by land
use and are conducted in small (~300 km²) watersheds, referred to as
management scale watersheds (Gellis and Walling, 2011). Sediment
fingerprinting also has been conducted in larger watersheds
(1000’s km²) (Collins et al., 1997b; Douglas et al., 2003; Voli et al., 2013) but at this larger scale, source percentages by sub-basin
and geologic provenance become more important. Here, a
confluence-based approach can be the most efficient means of
rationalising source sampling through the collection of sediment
samples upstream and downstream of major tributaries (e.g. Vale
et al., 2016). Gellis et al. 2016 (this issue) examined sediment
sources for a large region of the United States, the Midwest cornbelt
(648,239 km²), through sampling of fine-grained bed material in 98
wadeable streams (ranging in area from 6.8 to 5893 km²). Building
upon an approach using fallout radionuclides, 7Be and 210Pbex
(Matisoff et al., 2005). Gellis et al. (2016 this issue) estimated the
percentage of surface versus channel derived sediment, and the age
of this sediment to less than one year. Results indicate that the
majority of sediment is channel derived with many samples being
less than 100 days old.

A key outstanding issue is the optimisation of source sampling
strategies informed by an understanding of the variability of tracers
in the sources concerned (cf. McBratney and Webster, 1983;
Oldfield et al., 1989; Sutherland, 1991). Here, one fundamental
issue is the collection of sufficient sample numbers for statistical
robustness. Probability sampling designs (cf. Collins et al., 2001a)
have not been widely adopted, primarily because the implications
are that many more samples will need to be collected than is
currently normal practice and permitted by research budgets. In an
attempt to deal with such issues, many studies collect multiple sub-
samples within the immediate vicinity of a specific point and bulk
these into a composite representative of an individual sampling
5. Tracer selection for source discrimination

Following early studies that tended to pre-select tracer shortlists (e.g. Peart and Walling, 1988), most applications of sediment fingerprinting have measured multiple tracers in source samples and then applied statistical tests to confirm source discrimination. Here, once again, recent work has underscored the need to consider carefully a number of critical factors pertaining to either pre-selection, or further screening, following analyses of source material samples. Firstly, confirmation of a sound physical basis for any tracer providing discrimination between potential sources is highly preferable (Foster and Lees, 2000). Accordingly, and by way of example, prior geochemical knowledge linked to geological variation can be used to guide initial tracer selection (Laceby et al., 2015). Equally, an understanding of tracer environmental behaviour, such as that responsible for the contrasting fallout radionuclide signatures of surface and subsurface sources, can be used as a basis for selecting these particular tracers (Walling et al., 2003a). Prior knowledge of the impact of weathering processes in enriching or depleting tracers in specific sources (e.g. surface soil) can also be used (Koiter et al., 2013). Secondly, in the context of the potential for tracer perturbation, composite signatures should not necessarily be based on reductionist optimisation, since larger composite signatures can reduce uncertainty and help counter problems associated with the perturbation of any individual tracer (Sheriff et al., 2015). Here, there is a need to consider expanding composite signatures in the context of goodness-of-fit metrics for unmixing model performance. Thirdly, tracers with small differences between source groups should not be used since these generate larger uncertainties in estimated source proportions (Pulley et al., 2015a). Fourthly, tracers with greater between-group to within-group variability ratios should be pre-selected for inclusion in statistical tests applied for quantifying source discrimination (Pulley et al., 2015a). Previous work has shown that individual tracer property groups can provide robust discrimination (Collins and Walling, 2002), but where resources permit, the inclusion of properties responding to differing environmental controls is preferable. Although prior knowledge of tracer behaviour may not be for the precise physiographic setting in question, it is likely that sufficient general guidance on tracer pre-selection can be deduced from existing understanding of the typical environmental behaviour of most tracers.

6. Selection of grain size fractions for tracer analyses

The most common practice in published fingerprinting studies is to fingerprint the <63 μm fraction of sediment. The initial justification for this selection, above and beyond the dominant proportion of fluvial suspended sediment loads being represented by this size fraction, was to limit particle size effects given the knowledge that particle size exerts a strong influence on many of the tracers used for fingerprinting (e.g. Jonasson, 1977; Horowitz, 1991). But, because it has been shown that substantial variability in tracer concentrations can exist even within the <63 μm fraction (e.g. Horowitz and Elrick, 1987; Walling and Woodward, 1992; Motha et al., 2003; Hatfield and Maher, 2009; Pulley and Rowntree, 2016), an alternative approach is the use of narrower particle size fractions. Wallbrink (2004), for example, used only the <10 μm fraction, significantly reducing the capacity for particle size variability in the traced fraction. Hatfield and Maher (2009) found that the magnetic properties of catchment soils were significantly different between different particle size ranges of the same source material. As a result, they separated the sediment into 31–63, 8–31, 2–8 and <2 μm aliquots and the contribution of each fraction to the total magnetic properties of the sediment were quantified. Whilst
the methods of Hatfield and Maher (2009) have distinct advantages for identifying particle size effects and selecting the optimum particle size for tracing, they do require the measurement of tracers on multiple particle size fractions, significantly increasing the time and cost of analyses. Therefore, the selection of a narrow particle size range may be of benefit to many fingerprinting studies and, accordingly, some have pre-selected restricted ranges in their procedures (e.g. Wallbrink et al., 2003; Douglas et al., 2010; Wilkinson et al., 2013; Theuring et al., 2013, 2015; Lacey et al., 2015; Haddachi et al., 2015). Since finer fractions are more geochemically active, they are likely to provide more robust source discrimination; however this benefit can be counterbalanced because these finer grain size ranges also are more susceptible to transformation and non-conservative behaviour during transport. It also can be cost prohibitive to obtain sufficient sample masses of restricted size ranges to permit tracer quantification. Selection of any individual size fraction will only be appropriate if this is shown to be the size class that represents the majority of sediment in transport and indeed the fraction responsible for the environmental issue(s) (e.g. degradation of a coral reef or siltation of salmonid spawning gravel(s) in question. If sediment fingerprinting is to become a widely used management tool, the ability to source individual fractions and/or using very limited size ranges of fine-grained sediment may be cost prohibitive.

7. Tracer conservatism

Sediment source fingerprinting techniques are based on the fundamental assumption that selected tracer properties behave conservatively during mobilisation and delivery through the catchment system and that the properties of source material and sediment samples can therefore be directly compared. The significance of this assumption is increasingly recognised, but also challenged (Foster and Lees, 2000; Motha et al., 2002a,b, 2003, 2004; Koiter et al., 2013; Smith and Blake, 2014; Kraushaar et al., 2015). Early work highlighted the paucity of understanding on this critical assumption (Bubb and Lester, 1991; Zhang and Huang, 1993). Chemical transformations can occur in conjunction with a range of mechanisms throughout the sediment cycle, including, amongst others, scavenging by Fe/Mn oxides, chemical precipitation and incorporation into crystalline matrices (Forstner and Salomans, 1980; Foster and Lees, 2000). Despite these risks, published studies have included tracers prone to transformation, including phosphorus fractions (e.g. Owens et al., 2000). Whilst there are risks of non-conservative behaviour for actively transported fine-grained sediment, such risks are potentially elevated where sedimentary deposits are used to reconstruct catchment sediment sources through time. Post-depositional dissolution or diagenesis and the in-growth of bacterial magnetite can, for example, impact the conservatism of mineral-magnetic tracers (e.g. Foster et al., 2008; Pulley et al., 2015b). Short-lived radionuclides (e.g. 210Pb, 137Cs and 8Be) are also unsuitable for long-term (more than 100–150 years old) tracing as their activities will be influenced by fallout histories and short half lives. Longer-lived gamma emitting radionuclides such as 40K and 232U, will be more suitable for long-term (centuries to millennia) source reconstructions, assuming they provide robust source discrimination, because of their much longer half-lives (Walling and Foster, 2016).

Of the sources of uncertainty highlighted in the published literature, the effects of changing sediment particle size and organic matter content on tracer signatures during the sediment cycle through catchment systems are often prominent. The effects of these factors on many of the geochemical properties commonly used as sediment source tracers was recognised early on (e.g. Goldberg, 1954; Rex and Goldberg, 1958; Goldberg and Arrhenius, 1958; Krauskopf, 1956; Kononova, 1966; Jones and Bowser, 1978; see Horowitz, 1991 for additional references), yet little of such work seems to have been integrated into current source tracing procedures. Associations of many elements with organic material are often unpredictable, with some elements having a greater affinity than others (Swanson et al., 1966; Saxby, 1969; Rashid, 1974; Bunzl et al., 1976; Jonasson, 1977; Maulé and Dudas, 1988; see Horowitz, 1991 for additional references). The strength of these associations may differ between catchments (Gibbs, 1977) and organic matter can behave as both a diluent, (e.g. magnetic signatures (Walling and Foster, 2016)) or as a contributor (e.g., Horowitz, 1985; Horowitz and Elick, 1987). Organic corrections are widely used in conjunction with the application of mineral-magnetic fingerprints. Efforts to mitigate the effects of particle size and organic matter in fingerprinting studies can therefore be seen as being in an early stage of development with many investigations neglecting to include any significant attempt to mitigate their effects other than to sieve to <63 μm and employ elementary corrections, as discussed elsewhere in this paper.

In the absence of comprehensive information and guidance on the conservatism of multiple tracers in different environments, the vast majority of studies continue to use a simple screening technique to evaluate the conservative behaviour of various tracers based on the so-called range or bracket test, using a variety of rules (e.g. Foster and Lees, 2000; Wilkinson et al., 2013; Collins et al., 2013a,b,c; Gellis and Nee, 2013). A principal danger with existing range tests is that whilst they confirm that non-conservative transformation is not significant in the context of the sampled source tracer ranges, they do not confirm the complete absence of any non-conservatism (Collins et al., 2013b,c). The use of the range test can be underpinned by tracer screening using literature reviews dealing with tracer geochemistry in conjunction with an understanding of the various effects of changing physicochemical conditions between the source area and the sink (Kraushaar et al., 2015) and this pragmatic approach merits further attention. Pulley et al. (2015c) produced bi-plots of magnetic properties for source samples comprising different particle size fractions and were able to identify if lake sediment samples exhibited similar linear relationships, suggesting the conservatism of these tracers in the deposited sediment. This approach represents a more robust form of the range test, although it does greatly increase the time and cost requirements for tracer analysis and is dependent on a relationship between at least two tracer variables.

Apportionment modelling in the procedures used by some researchers (e.g. Motha et al., 2004) has attempted to include the impact of non-conservative tracers explicitly. Here, for example, work by Collins et al. (2010b, 2012a,b, 2013b,c, 2014) has used probability density functions (pdfs) to construct deivate target sediment tracer values which are then sampled during un-mixing model Monte Carlo repeat iterations using a Latin Hypercube. This approach recognises explicitly that any individual sediment sample, or the sediment from any individual location in a catchment system, has the potential to be transformed due to selectivity and/or biogeochemical alteration (e.g. sorption, dissolution, precipitation, oxidation, reduction), but that collectively, those samples will provide a range of more and less altered tracer values which can be treated as a ‘conservative’ population (conservative in the context of using the simple range test). Sheriff et al. (2015) report the use of a tracer permutation algorithm developed by Franks and Rowan (2000) to determine the impact of non-conservative tracers on source apportionment predictions. The accuracy of predicted mean source contributions was reported to be significantly different between the maximum positive and negative levels of tracer corruption (-90 and +15%), but uncertainty was not impacted by mimicking tracer transformation.
8. Source apportionment modelling

Use of mathematical techniques (e.g. Yu and Oldfield, 1989, 1993; Walling et al., 1993; Walling and Woodward, 1995; Collins et al., 1997a; Gellis and Landwehr, 2006; Hughes et al., 2009; Sheriff et al., 2015) to un-mix sediment samples represents a key methodological component of source fingerprinting procedures over the past two decades. Recent studies using un-mixing models have applied different composite signatures to assess variation in predictions dependent on the tracers used and to improve the use of multiple tracers provided by current analytical techniques including ICP-MS (Collins et al., 2012a, 2013c; Stone et al., 2014; Theuring et al., 2015) and NIRS (Collins et al., 2013b, 2014). The explicit assessment of uncertainty in conjunction with the growing application of un-mixing modelling was first introduced by Franks and Rowan (2000) and Rowan et al. (2000) in the form of Monte Carlo analysis. It is now standard to include an explicit assessment of uncertainties in conjunction with the use of source apportionment modelling.

The growing use of sediment un-mixing models has demonstrated that the range of uncertainty outputs from Monte Carlo routines is primarily driven by the within-source group variability in tracer concentrations and the corresponding differences in tracer concentrations between-source groups (Small et al., 2002). As a result, weightings have been applied to give a larger emphasis during un-mixing modelling to tracers with a lower within-source variability and greater discriminatory power (Martinez-Carreras et al., 2008; Collins et al., 2010c; Wilkinson et al., 2013; Gellis and Noe, 2013). The latter weighting has, in some cases, been used as a substitute for original weightings reflecting the analytical errors or precision associated with individual tracers (Mackas et al., 1987; He and Owens, 1995; Collins et al., 1997a). These weightings were developed in response to some papers identifying the need to explore their use (Walling et al., 1993). All such weightings should be carefully assessed in the context of evaluating modelled source proportions using goodness-of-fit metrics and artificial sediment mixtures (e.g. Lacey and Olley, 2015). The sensitivity of modelled source proportions to these types of weightings has been reported as limited based on some datasets (Pulley et al., 2015b) and where the impacts are greater, the weightings reflecting analytical precision or tracer discriminatory power are subtly compared to other weightings (Lacey and Olley, 2015).

More recently, variability ratios (of inter-/intra-source group variability) have been recommended by Pulley et al. (2015a) to capture the fundamental need to select tracers that maximize between- rather than within-group tracer variation. These variability ratios can be applied as an initial screen in the tracer selection procedure to remove tracers that are likely to result in elevated levels of uncertainty in both source discrimination (Collins and Walling, 2002) and un-mixing model outputs (Pulley et al., 2015a). Some work has also introduced distribution-based modelling, to ensure that multiple model iterations for uncertainty analyses maintain relationships between tracers during the iterative sampling of tracer distributions reducing the uncertainty ranges present in model outputs (Lacey and Olley, 2015; Lacey et al., 2015). In terms of the input tracer distributions, a critical decision is whether to represent source groups using the 25th–75th percentile range or the 5th–95th percentile range since this decision alone can influence the corresponding uncertainty ranges associated with modelled source proportions. Regardless of the scaling used, mixing model outputs are characterised by uncertainty ranges and a key decision is how to present this uncertainty to catchment stakeholders. Here, many existing studies have reported gross uncertainty ranges (e.g. 5th–95th percentiles, or the entire pdf), the average mean or median source proportions with associated uncertainty (95% confidence limits) and tested the convergence of the model runs (e.g. Collins et al., 2013c). Communicating the uncertainty ranges to stakeholders involved in decision-making for managing the sediment problem is important. To simplify the communication of uncertainty, whilst taking explicit account of this issue, some researchers have calculated relative frequency-weighted average mean or median source contributions (e.g. Collins et al., 2013b,c; 2014). The processing of sediment source tracing data for a single location within a study catchment will always be prone to bias introduced by the scale dependencies associated with spatial variation in the mixtures of potential sources and corresponding geomorphic processes driving sediment mobilisation and delivery.

The adoption of un-mixing models by many studies has been accompanied by the inclusion of particle size corrections. In the majority of studies, these continue to be based on the assumption of a simple linear relationship between particle size and tracer signatures (e.g. Collins et al., 1997a; Owens et al., 1999, 2000; Walling et al., 1999, 2003a, 2006, 2008; Smith et al., 2011; Smith and Blake, 2014). However, it has been recognised that relationships between particle size and many tracer signatures are non-linear, especially for specific surface areas >1.0 m² g⁻¹ (Horowitz and Elrick, 1987; He and Walling, 1996; Foster et al., 1998; Russell et al., 2001; Motha et al., 2003; Bihari and Dezs, 2008; Hatfield and Maher, 2009; Oldfield et al., 2009); thus introducing uncertainties in conjunction with simple linear corrections. Previous work has demonstrated that significant contrasts can exist in particle size composition between different source and sediment samples, even when all samples have been screened through a 63 µm sieve (e.g. Walling and Woodward, 1992; Russell et al., 2001). Such data imply that even post sieving to <63 µm, the tracers of source and sediment samples cannot be directly compared without further correction. Motha et al. (2003) measured tracer signatures associated with various particle size fractions and developed tracer signature-specific correction factors. Russell et al. (2001) also developed tracer-specific curvilinear corrections rather than assuming a generic linear relationship between concentration and grain size. Whilst such approaches help mitigate uncertainties associated with linear corrections they do, however, have disadvantages in terms of the time required for laboratory work. Due to these challenges, some studies have used enrichment factors based on the measured concentrations of tracers in sediment and source samples (e.g. Peart and Walling, 1986; He and Owens, 1995). Alternatively, other studies have adjusted source material tracer concentrations by using information on the grain size of target sediment and tracer concentrations associated with particle size fractions of source materials, to estimate property concentrations in source material with the same grain size composition as the target sediment (Walling and Woodward, 1992; Slattery et al., 1995; Motha et al., 2004). Recognizing that the relation of grain size and tracer property can be positive, negative, or have no relation, Gellis and Noe (2013) used regression analysis of the D₅₀ of source samples against tracer concentration to produce a grain-size correction factor. This has the advantage that the fractionation of source samples and analysis of each fraction is not required and a linear relationship is not assumed. However, such methods may require extrapolation of a trend line beyond the range of values found in the source samples, thereby introducing uncertainty. An alternative to developing corrections for grain size effects, is to use narrower size fractions (e.g. <10 µm) to counter the potential influence of selectivity during the sediment delivery cascade (e.g. Theuring et al., 2015).

Elementary organic matter corrections have also been used (e.g. Walling and Woodward, 1982, 1986; Collins et al., 1997a; Motha et al., 2003, 2004; Walling et al., 2003a, 2006, 2008; Gellis and Noe,
2013; Pulley et al., 2015b), driven primarily by correlations between tracer concentration and organic matter content and by the improvements in the goodness-of-fit outcomes for un-mixing. Again, these weightings were often developed in response to some researchers identifying the need for their inclusion (Walling et al., 1993), although such correlations are site-specific, meaning there is no universal correction factor. Although some research has highlighted the risk of such corrections biasing source predictions (Smith and Blake, 2014); alternatively, recent research has shown that they have limited impact on the source estimates (Pulley et al., 2015b). The importance of carefully assessing elementary corrections for grain size and organic matter on a dataset-specific basis and making informed decisions to avoid over-correction has long been underscored (e.g. Walling and Collins, 2000; Walling et al., 2003a). Noise associated with differing organic matter contents of source materials, or a high within-source group variability in tracer concentrations may mask relationships between DS and tracers leaving them unaccounted for. Sediment-associated organic matter in the fluvial environment exists as loosely-bound particulate material (e.g. leaf litter), which in the case of many of the most commonly used tracers (apart from biomarkers) will act as a diluent, and as surface coatings for mineral particles, where it can act as a concentrator. Moving forward, this implies that the development of more informed correction factors for organic matter needs to take into account both grain-size and phase specific aspects of the problem. Such work has important resource implications.

Some research during recent years has been directed towards the comparison of variations in source apportionment depending on the applied un-mixing model. Haddadchi et al. (2014), for example, compared four different model structures using artificial mixtures with known proportions of sediment sources. There has also been a growing number of source tracing studies which, rather than using maximum likelihood/frequentist (see modelling papers cited above) methods, instead, use Bayesian (Fox and Papanicolaou, 2008; Rowan et al., 2011; Massoudieh et al., 2012; D’Haen et al., 2013; Cooper et al., 2014) modelling approaches. Uptake of the Bayesian models has benefitted from some of them (e.g. Barthod et al., 2015) being open source. The need to compare local and global solutions using the former types of models has been underscored by previous work (Collins et al., 2010d). Model structure and the robustness of the input data both have a strong bearing on the outputs, and end users must carefully assess model structures and approaches when applying fingerprinting procedures. Numerous uncertainties which are not fully accounted for in current fingerprinting procedures have been identified in recent publications, highlighting the need for further methodological refinements which, where appropriate and underpinned by replicated evidence based on multiple catchments and environmental settings, need to be incorporated into sediment un-mixing models (e.g. Motha et al., 2002a,b; D’Haen et al., 2012; Koiter et al., 2013; Smith and Blake, 2014; Pulley et al., 2015b; Lacey and Olley, 2015). An ongoing problem is that many papers assess specific issues for a single or limited set of study catchments/environments, and then propose generic guidance which simply may not be widely applicable. Importantly, however, these recent studies serve as useful reminders that source tracing datasets should be treated on a case by case basis.

9. Use of artificial sediment mixtures to assess source apportionment modelling

The use of artificial mixtures of known quantities of sediment sources (cf. Stott, 1986) has gained increasing popularity in recent years and represents an important component for the development of robust, widely applicable source tracing procedures. A limitation of fingerprinting research is that it is difficult to validate estimated source proportions using independent evidence as the monitoring and measurement techniques required face their own limitations in terms of the practicalities and costs of deployment both spatially and temporally (Collins and Walling, 2004). Validation of fingerprinting estimates against independently measured data assembled using alternative techniques therefore continues to be rare, although some examples exist (e.g. Peart and Walling, 1988; Collins et al., 1998; Stone et al., 2014). Mixtures of known proportions of sediment sources thereby provide a pragmatic opportunity to assess the accuracy of a fingerprinting procedure on the basis of its estimated source proportions. Early studies using artificial mixtures include the work by Lees (1997) who identified non-linear additivity associated with the use of the mineral magnetic properties of sediment. Franks and Rowan (2000) used five artificial mixtures consisting of five source types based on major chemical groups to assess a source tracing procedure. Small et al. (2004) used a Bayesian modelling approach and artificial mixtures to explore sample sourcing related uncertainties and the number of source samples required to limit uncertainty in modelling results. Additional studies using artificial mixtures to assess un-mixing model outputs include those by Hughes et al. (2009), Poulenard et al. (2012), Legout et al. (2013), Brosinsky et al. (2014), Haddadchi et al. (2014) and Lacey and Olley (2015). Given the laboratory work associated with generating and analysing the tracer content of artificial source mixtures, some recent studies have introduced synthetic mixtures based on Monte Carlo routines (Palazón et al., 2015; Sheriff et al., 2015) as an alternative. Whilst virtual sample mixtures can be deliberately corrupted to mimic uncertainty (Sheriff et al., 2015) they do, however, have limitations including, for example, different source groups having contrasting particle size distributions (Palazón et al., 2015).

10. A decision tree for guiding application of sediment source tracing

Progress continues to be made in the refinement of sediment source fingerprinting procedures but much scientific debate is ongoing. Following four decades involving preliminary applications, acceptance of the need for composite signatures and the introduction of statistical and numerical modelling approaches, including uncertainty assessment, recent work has re-visited critical assumptions and challenged some recent proposed methodological modifications. In the context of ongoing studies, and the diverging opinions on some aspects of fingerprinting procedures, it is timely to propose a revised decision-tree for supporting critical choices that have to be made by end-users applying the technique. This decision-tree (Fig. 1) attempts to capture the current state-of-the-art, and hopefully serves as one means of synthesizing the lessons gleaned from the past 40 years of research.

Currently many steps of the methodology presented in the decision-tree are in the early stages of research and firm instruction cannot be given due to many factors or processes being site-specific. However, the decision-tree aims to provide an overarching comprehensive methodology which includes important steps for evaluation, validation and uncertainty analysis. It is intended that the decision-tree will provide a framework from which researchers and reviewers can structure their methods and interpretation(s) of sediment fingerprinting results. The goals and resource availability of different studies will likely mean that not all stages of this decision-tree can be strictly followed but in such situations, end-users must identify limitations and shortcomings in the procedures actually applied when reporting their results.
Sediment and source sampling

Is historically-deposited sediment being traced?

<table>
<thead>
<tr>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>The method of sampling actively transported sediment will be dictated by the aims of the investigation and the need to consider both temporal and spatial representativeness. Instantaneous channel bed sediment sampling using bed disturbance (see Lambert and Walling, 1988; Duerdoth et al., 2015), provides a useful means of collecting a large number of replicate sediment samples both temporally and spatially. Alternatively, passive time-integrated interstitial (Walling et al., 2003a) or suspended sediment (Phillips et al., 2000) traps provide alternative approaches for improving temporal representativeness. Replicate sediment samples should be collected to account for the uncertainty associated with the sediment sampling methods. Multiple channel sites should be covered to take into account scale dependency and process domains. Based on the type of tracers being analysed, care should be taken during sampling not to contaminate the sample; for example, restrict using metal samplers if metals are among the tracers to be analysed. Consideration should also be given to the sediment mass required for tracer analyses. Fractionation to a narrow particle size range may be required, reducing the mass of sample available for analyses. (Note: some analyses are non-destructive such as gamma spectrometry and mineral magnetic analysis and these analyses should be undertaken first if sample size is an issue).</td>
<td></td>
</tr>
<tr>
<td>When investigating the sources of historically-deposited (e.g. lake, reservoir, floodplain, wetland, estuary) sediment, special consideration should be given to potential factors impacting the sediment and its associated tracers. A lake that regularly dries out may be subject to the reworking of sediments as channels of water flow across the lake bed. Additionally, lakes which are used by large animals may have their sediment deposits disturbed, disrupting the down-core chronology. Very waterlogged floodplains or wetlands are likely to store sediment under anoxic reducing conditions and a high organic matter content will also contribute to this problem, resulting in the dissolution of tracers. A grey mottled gleyed appearance of the sediment is a good indication that tracer dissolution is severe. If this is the case, source tracing is unlikely to be successful. Instead, consider sampling a part of the floodplain where sediment is stored in drier more chemically-stable conditions. Alternatively, rather than sampling a wetland, if there is an alluvial fan delivering sediment into to it, then sampling the fan may yield more reliable results. Replicate cores or surface scrapes should be collected from the receptor rather than using a single-core or sample approach to assist inclusion of uncertainty for target sediment signatures.</td>
<td></td>
</tr>
</tbody>
</table>

The number of source samples collected will depend on the size and homogeneity of the study catchment. The aim of the source sampling campaign should be to capture adequately the variability of each source group. It is important to be flexible in how source groups are defined at this stage in the procedure. The geology, land use and soil type of the study catchment should all be considered as source groups and it should be ensured that an adequate number of source samples are collected from each of the different source categories, potentially informed by probability sampling. In the context of resource constraints for analytical costs, replicate sub-samples at each location selected for each source should be bulked into composites. However, a sufficient number of individual samples must still be collected so that variability can be sufficiently represented in the statistical analyses and modelling. The results derived by Small et al. (2002) suggest that fewer than 20 samples in a source group is likely to result in a high amount of uncertainty associated with apportionment modelling results. However, where a large within-source group variability exists, a greater number of source samples may be required. For surface sources, sample to the typical depth that sediment generation processes operate at (e.g. 0-2 cm depth has been widely used in temperate environments; Walling and Woodward, 1995). Deeper sampling may be appropriate in arid or semi-arid landscapes.

Fig. 1. A methodological decision-tree for guiding application of sediment source fingerprinting (Rousseeuw and Croux, 1993).
Tracer selection

How large is your budget?

Large budget

Are you tracing historically-deposited sediment?

Yes

Sediment colour has been shown to be an effective tracer in arid and semi-arid environments where sediments are present with a low organic matter content. However, colour may be made non-conservative by the organic coating of sediment particles in productive rivers and lakes.

No

Gamma spectroscopy will permit the use of $^{137}$Cs and $^{210}$Pb which are effective discriminators on the basis of land use and subsurface sources. Lithogenic radionuclides are also measured by this method (at no extra cost and little extra effort) and can be effective discriminators of different geologies.

Medium budget

Sediment colour has been shown to be an effective tracer in arid and semi-arid environments where sediments are present with a low organic matter content. However, colour may be made non-conservative by the organic coating of sediment particles in productive rivers and lakes.

No

Gamma spectroscopy will permit the use of $^{137}$Cs and $^{210}$Pb which are effective discriminators on the basis of land use and subsurface sources. Lithogenic radionuclides are also measured by this method (at no extra cost and little extra effort) and can be effective discriminators of different geologies.

Small budget

Are you tracing historically-deposited sediment?

Yes

Does your catchment contain heterogeneous geology?

Yes

Chalk and limestone geologies are likely to be rich in calcium and igneous or ironstone rock types are likely to be more magnetic or have higher iron concentrations than sedimentary rock types. Magnetism, geochemistry and bulk stable isotopes are likely to be effective tracers in many catchments.

No

Mineral-magnetic signatures have been shown to discriminate between surface and subsurface sources. However, this discrimination may be caused by the dissolution of magnetic grains in anoxic subsurface conditions. Therefore, care should be taken when tracing historically-deposited sediment to ensure that similar dissolution is not occurring in the sediment receptor, destroying the basis for source discrimination. Magnetic signatures are often correlated with each other which is of benefit when performing a mass conservation test but may limit source discrimination. They can also easily be corrected for organic matter content.

Urban road dusts and damaged road verges may be expected to have different magnetic properties than cultivated and grassland. Similarly, they are expected to have higher concentrations of geochemical tracers such as lead and zinc.

Geochemical tracers may also discriminate on the basis of land use, but the basis for discrimination is less well established than for magnetic tracers. For example, differences between source groups may reflect a purely statistical solution, which is presumed to be caused by geochemical differences due to factors such as weathering or anthropogenic applications.
Tracer conservation must be assessed using a range or bracket test as a bare minimum. In addition, the tracer variability ratio can be used for screening tracers, as can expert opinion and prior experience on the general conservatism of individual tracers in specific environments. Additionally, when tracing historically-deposited sediment, any tracer that is likely to have changed in concentration in the source groups over the time period of sediment deposition should be removed e.g. lead from vehicle emissions or phosphorus due to artificial fertilizer applications.

**Organic matter**

Do your sediment or source samples generally have an organic matter content above 30%? (this value is based on the authors’ experience to date and therefore should be treated as preliminary guidance. Further research into organic matter – tracer relationships is clearly required)

- Yes
  - Is the tracing of the sources of sediment-associated organic matter one of your research objectives?
  - Yes
    - Trace the sources of sediment-associated organic matter (e.g. Collins et al., 2014)
  - Yes
    - Consider the use of mineral-magnetic signatures which can easily be corrected for organic matter content with a simple data correction. Alternatively, consider the use of a coarse particle size fraction e.g. 63 - 32 μm, where organic matter can be poured off during the wet sieving process and a small surface area is available for the coating of sediment particles by tightly-bound organic matter.
  - No
    - Organic matter is unlikely to be a major source of uncertainty in your investigation.

- No
  - Do your sediment or source samples generally have an organic matter content above 5%? (see comment above)

*Fig. 1. (continued).*
Although the steps in this new decision-tree may increase the resource requirements compared with those used by past published studies, it is our view that adequate recognition must be given to the current state-of-the-art for sediment source fingerprinting by providing appropriate caveats and/or reporting levels of uncertainty, especially where the results are intended to inform catchment management and sediment mitigation planning.
Notes

It is likely that sediment source discrimination will be in some way particle size dependant in many catchments. For example, sedimentary limestones, chalks and ironstones are likely to have the majority of their calcium and iron concentrated in very fine grain sizes, meaning that discrimination could be weaker in the coarse silt and sand fractions than in fine silts. Additionally, weathering and soil formation processes are likely to result in the preferential precipitation of weathering products on the large surface areas of fine silt and clay particles; as might the adsorption of fallout radionuclides take place. Therefore, prior knowledge of the controls on source group tracer concentrations combined with objective particle size selection may be useful for achieving optimal source discrimination.

The potential for particle size related uncertainties in a tracing methodology is likely to be proportional to the range of particle size selected for analysis. For example, there is far less scope for particle size change in the <10 μm fraction than in the <63 μm fraction. Therefore, when using wide particle size ranges, it is important to demonstrate that particle size effects have been properly investigated and, where necessary, accounted for. The result validation section provides guidance on the use of artificial mixtures to demonstrate the range of uncertainty potentially caused by particle size effects. It is also good practice to compare the particle size distribution of the prepared source samples to the target sediment samples collected from the lake, floodplain or river to show if there are significant differences between the two.

Fig. 1. (continued).
**Particle size and organic matter corrections**

Notes: if using a suitably narrow particle size range this stage may not be necessary. However, it should be considered if time and resources allow.

Testing the relation (i.e. regression analysis) of grain size to each tracer’s concentration (or activity) may be useful in determining whether a size correction should be applied.

Was a narrow range of particle sizes analysed?

- **Yes**
  - Develop a specific correction factor for each source group and each tracer using the methods of Motha et al. (2003).

- **No**
  - When producing a scatter plot of particle size ($D_{50}$ or specific surface area) or organic matter content against tracer concentration for each source group, are there any significant relationships (Gellis et al., 2015)?
    - **No**
      - Do not use any corrections but ensure that the particle size range being traced is sufficiently narrow to limit error caused by particle size effects.
    - **Yes**
      - Use these relationships to form correction factors for each target sediment sample.

The conservatism and discrimination tests must be performed after any correction factors have been applied to the data as the basis for source discrimination is very likely to be different for different particle size fractions. Due to particle size related differences in the basis for source discrimination, these correction factors may not always be effective unless source discrimination is fully incorporated into the correction methodology.

The effectiveness of any developed corrections must be evaluated during the "methodology validation" stage of the procedure. If they do not improve the accuracy of the tracing they should not be used.

**Fig. 1.** (continued).
Source group classification

Does the study catchment have homogenous geology?

- Yes
  - Adopt a standard source classification for the study goals e.g. by land use. But strongly consider trying the cluster analysis based methodology.

- No
  - Consider using a cluster analysis source classification method based upon the methods of Walling et al. (1993). Map the source samples according to cluster group on a map of the catchment and decide upon what each cluster represents e.g. a specific geology.

Do the resulting cluster derived source groups suit the study aims?

- Yes
  - Reclassify the cluster analysis groups to suit management goals e.g. into a surface and subsurface components or a cultivated and grassland component. See Pulley et al. (2016).

- No
  - The range of uncertainty in your final modelling results is likely to be unacceptably large. Consider a different source group classification scheme, e.g. a different number of clusters or different modification of cluster groups.

Calculate tracer variability ratios for each pair of source groups for each tracer using the formula; 

\[
\frac{((\text{maximum mean tracer concentration in source group} - \text{minimum mean tracer concentration in source group})}{\text{minimum mean tracer concentration in either source group})}{\text{the mean coefficient of variation (%) of the pair of source groups.}}
\]

Are the ratios generally greater than 1.0 meaning that inter-group variability is greater than intra-group variability? (Pulley et al., 2015a)

- Yes
  - The source groups are likely to be acceptable. However, consider testing if an alternative classification scheme may improve the variability ratios.
    - Any tracer with a maximum variability ratio lower than 1.0 should be removed from the procedure at this point.
    - The threshold of 1.0 can be increased as higher ratios will result in less uncertainty in the final results.

- No
  - The range of uncertainty in your final modelling results is likely to be unacceptably large. Consider a different source group classification scheme, e.g. a different number of clusters or different modification of cluster groups.

A smaller number of sediment source groups has been shown to potentially reduce uncertainty in modelling outputs. However, fewer groups with a much higher within-source group variability is also likely to increase uncertainty, so a balance is required. The tracer variability ratio can be used to determine when decreasing the number of source groups results in a large increase in within-source group variability.

Outliers and misclassified samples

At this point it is appropriate to examine the tracer signatures in the source groups to identify any extreme outlying values or if any source sample is likely to have been misclassified. Outliers may be defined such as greater or less than 3-times the standard deviation of the mean. Such outliers may introduce greater uncertainty into the modelling outputs. It will be a matter of personal judgement as to which source samples to remove. Obvious outliers should be removed but taking care to maintain the proper range of variability in the source group samples.

Fig. 1. (continued).
Conservatism tests

Consider shortlisting tracers using published evidence or expert opinion on conservatism

Are you tracing historically-deposited sediment?

Yes

Is there a significant down-core reduction in tracer signatures or a sudden loss of a tracer below a certain depth? See below for example;

Reduction

Yes

It is possible that the tracers are undergoing dissolution diagenesis. If the point at which dissolution takes effect is easily identifiable e.g. at 40 cm depth in the above figure, then remove all samples below that depth from further analysis or find a tracer which you can be sure is resistant to dissolution.

Are any of the measured tracers significantly correlated with each other?

No

Fig. 1. (continued).
Note:
The use of mineral magnetic signatures with this test may prove sensitive to the dissolution of ultrafine super paramagnetic (diameter <0.02 μm) or stable single domain (0.4 - 0.02 μm) grains. Should these grains be conservative it is unlikely that the dissolution of iron oxides and their associated geochemical and magnetic tracers is taking place.

Yes

For all tracers not significantly correlated with another tracer use the below. Those passing the above test can be used in the determination of the composite fingerprints stage of the methodology.

No

Use a conventional range test. Test with the 0th - 100th and/or 25th - 75th percentile ranges of the source groups to determine if the tracers in the sediment samples fall within the maximum and minimum values found in any source group (e.g. Collins et al., 2013c). Repeat using the mean or median values for source and target sediment samples (e.g. Wilkinson et al., 2013).

Most tracers pass the 25th - 75th percentile test for most sediment samples (>95%) and the means/medians range test.

Most tracers pass the 0th - 100th percentile test for most sediment samples (>95%) and the means/medians range test.

Most tracers do not pass the range tests.

Only take these tracers through to the next stage of the methodology. Be aware that your source groups may not be a good fit to the tracer signatures or there may be some form of tracer non-conservatism not detected by the range tests.

Only take these tracers through to the next stage of the methodology. Be aware that your source groups may not be a good fit to the tracer signatures or there may be some form of tracer non-conservatism not detected by the range tests.

Revisit your source group classification method or consider if there is a significant difference between the particle size or organic matter content of your source and target sediment samples. If tracing historically-deposited sediment significant dissolution may be taking place.

Fig. 1. (continued).
Source discrimination

Consider the use of a number of independent statistical tests to identify multiple composite signatures for discriminating the study catchment sediment sources. Consider the use of the Kruskal-Wallis H test, Principal Component Analysis, cluster analysis and linear discriminant analysis, amongst others. Consider the use of a two-step process with step one testing the ability of each individual tracer to provide some degree of discrimination between your source sources using the independent tests such as a Kruskal Wallis H-test. Consider ranking the results from each test to select the most powerful individual tracers (e.g. Collins et al., 2012a).

Step two takes the ranked results from each independent test and passes them through Discriminant Function Analysis driven by a stepwise algorithm to finalise the optimum composite signature. Consider driving the discriminant function analysis using a genetic algorithm.

Different optimum composite fingerprints

As we have few ways to validate the outcomes of a fingerprinting study, the replication of the modelling using multiple composite fingerprints comprising different sets of tracers is an important part of the procedure.

Larger composite fingerprints of tracers have been shown to reduce uncertainty in modelling outputs tested by the use of artificial sediment mixtures. However, this must be balanced with the need to minimise mixing model errors represented by the difference between source-weighted and measured sediment tracer values since larger fingerprints will return greater errors using a goodness-of-fit based on absolute error.

Fig. 1. (continued).
Source apportionment

The choice of mixing model structure will influence the outputs generated and the appropriateness of different structures should be explored. Consider the combined use of frequentist and Bayesian approaches. Models must include some form of Monte Carlo based uncertainty analysis to capture uncertainty in characterising the source and target sediment tracer values. Models must include the distributions of tracer signatures in both source groups and target sediment.

If sufficient composite samples (at least 20; Small et al., 2002) are collected per source group and/or for target sediment, use the distributions of the measured tracer values to construct pdfs for the apportionment modelling. If fewer composite samples are collected, Normality tests should be used to establish the most appropriate location (mean / median) and scale (standard deviation, median absolute deviation, Qn, Sn; Rousseeu and Croux, 1993) estimators for constructing the source and sediment tracer pdfs.

Using a model which maintains correlations between tracers in each source group can reduce the range of uncertainty in its outputs.

Run the un-mixing model, using an error threshold (e.g. disregarding iterations with an error >15%; Walling and Collins, 2000) to predict pdfs of source proportions. Use these to establish full uncertainty ranges. Test the reproducibility (convergence) of the model solutions by repeating the Monte Carlo analysis. Consider expressing uncertainty using relative frequency-weighted average mean or median source proportions (Collins et al., 2012a).

Estimate 95% confidence limits for these average means or medians.
Assess the goodness-of-fit between source-weighted and measured sediment tracers using a combination of absolute mean relative error (AMRE; Collins et al., 1997a) and mean relative error squared (MRES; Motha et al., 2003). Assess the relationship between these two estimators of goodness-of-fit for measured tracer values. Divergence between the two estimators is possible, especially with larger composite signatures. Acceptable results using these goodness-of-fit tests still need to interpreted alongside those under 'apportionment validation' using artificial mixtures.

Check the consistency of your source apportionment predictions using your different optimum composite signatures. Are the predictions based on each signature consistent?
To test the robustness of your model, put the source samples in as target samples and see how accurately they are ascribed.

Consider generating final source apportionment estimates by weighting the model results generated using each independent composite signature on the basis of a weighting combining the discriminatory efficiency of the signature and the associated AMRE.

Do the final source apportionment estimates make environmental sense for your study catchment?

Weightings may be included based upon within-source group variability or discriminatory power (Martinez-Carreras et al., 2008; Collins et al., 2010c) or tracer variability ratios (Pulley et al., 2015a). However, these may have a detrimental effect on model accuracy and should be tested using artificial mixtures of source groups before inclusion in the final methodology.

Apportionment validation

Prior to running the un-mixing model, methodological validation should be performed using the artificial mixing of known quantities of the sediment source groups. The mixtures should be used to validate the following:

Fig. 1. (continued).
Whether weightings for within-source variability and discriminatory efficiency impact on the accuracy of model outputs.

Whether the un-mixing model used provides accurate results and if the use of an alternative model can improve the accuracy of model results and decrease the full uncertainty ranges.

Whether differences in organic matter content between the source and target sediment samples is likely to be a significant source of error. Organic matter may be added to the mixture to judge the magnitude of the error likely caused by organic enrichment.

If correction factors for particle size and organic matter are used they should also be validated using the artificial mixtures.

Mixtures using only a small proportion of random sediment samples from each source group should also be used to determine how robust the methodology used is for sediment delivery from only a small spatial area of each source group. The reclassification of source groups may assist in reducing this particular source of uncertainty. Virtual sample mixtures can provide a time efficient means of completing this test, as can running a range of the source samples through the un-mixing model.

Many users of fingerprinting methodologies will decide not to fractionate samples to a very narrow particle size range e.g. <10 μm. When using wide particle size ranges, it is important to demonstrate that particle size is not a large cause of error in the results. This can be assessed using artificial mixtures sieved / settled to conform to the finest and coarsest sediment samples that are being traced. If this degree of precision is not practical timed settling may be used to roughly fractionate the mixtures into coarse and fine material which can then be run through the tracing methodology. This will provide an indication of the size of errors which could be caused by particle size differences.

Fig. 1. (continued).
Acknowledgements

The authors thank their co-workers on sediment fingerprinting studies and the participants in the workshop held at the 26th International Union of Geodesy and Geophysics meeting in Prague in 2015. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References


temporal sediment source dynamics in tropical ephemeral river systems. J. Soils Sediments 16, 1130–1144.


Theuring, P., Rode, M., Behrens, S., Kirchner, G., Jha, A., 2013. Identification of fluvial sediment sources in the Kharwa River catchment, Northern Mongolia. Hydrolog.


