Tracing spatial sources of suspended sediment in the Ohio River basin, USA, using water quality data from the NASQAN programme

YUSHENG ZHANG[1], ADRIAN L. COLLINS[1] & ARTHUR J. HOROWITZ[2]

1 *Environment Group, ADAS, Woodthorne, Wergs Road, Wolverhampton WV6 8TQ, UK*
yusheng.zhang@adas.co.uk

2 *US Geological Survey, Georgia Water Science Center, 3039 Amwiler Road, Atlanta, Georgia 30360, USA*

Abstract Documenting catchment suspended sediment sources remains critical for guiding the design of sediment management strategies and for abating the numerous environmental issues associated with enhanced loadings. Sediment fingerprinting techniques have received increasing attention in this respect since, at the catchment scale, they avoid many of the problems and uncertainties experienced with using more traditional measurement methods. As part of the US Geological Survey's revised NASQAN (National Stream Quality Accounting Network) programme, routine water quality samples are collected in selected large river basins in the United States. The geochemical data generated from these samples over a period of 10 years (1996–2006), were used as the basis of a fingerprinting exercise to assess the key tributary sub-catchment spatial sources of suspended sediment transported by the Ohio River. A Monte Carlo approach was used during the fingerprinting mass balance modelling to quantify uncertainty in the spatial source estimates. The results should be interpreted with respect to the routine but infrequent nature of the suspended sediment samples used as the basis for the sourcing exercise, but nonetheless, demonstrate how routine monitoring samples can be used to provide some preliminary information on sediment provenance in large drainage basins.

Key words sediment sources; fingerprinting; uncertainty; routine water quality sampling

INTRODUCTION

An understanding of the spatial and temporal dynamics of suspended sediment sources within a catchment is critical for supporting cost-effective and integrated management of sediment-associated environmental problems (Collins *et al*., 2001). Sediment has been identified as a significant vector of catchment diffuse pollution, controlling the transfer and fate of nutrients and contaminants (Catt *et al*., 1998; Warren *et al*., 2003), and is responsible for degrading freshwater habitats and detrimentally impacting aquatic biota (Packman & Mackay, 2003; Greig *et al*., 2005). It is therefore important to control the sediment problem at source, since prevention is better than cure. To meet this demand for reliable information on catchment sediment sources, sediment fingerprinting techniques have received increasing attention since they avoid many of the problems and uncertainties experienced with using more traditional measurement or monitoring techniques (Collins & Walling, 2004).

Established sediment fingerprinting procedures usually involve a number of key stages. These include, amongst others, the categorization of potential sources, targeted sampling of those source areas, selection of end-member sampling sites for determining source contributions, sediment sampling, laboratory analyses of fingerprint properties for source discrimination and, sediment source apportionment using a numerical mixing model. In many instances, sediment source fingerprinting has been used to apportion the relative contributions from individual sediment source types classified on the basis of land use and eroding channel banks (Collins *et al*., 1997; Russell *et al*., 2001; Wallbrink *et al*., 2003; Motha *et al*., 2004; Walling, 2005; Walling *et al*., 2008; Hatfield & Maher, 2009; Wilkinson *et al*., 2009). Alternatively, this approach has been used to apportion sediment contributions from spatial sources categorized on the basis of geological units (Collins *et al*., 1998; Walling *et al*., 1999) or tributary sub-catchments (Collins *et al*., 1996, 2009) and these alternative applications have proven especially useful for rationalising the interpretation of sediment source information in larger drainage basins with complex land use patterns. In the case of investigating spatial sediment inputs from tributary sub-catchments, suspended sediment samples are collected from the individual tributaries thereby substituting the collection of soil samples used in applications focusing on specific source types. The sets of tributary sediment samples are compared with corresponding samples collected at the overall catchment outlet, using fingerprint properties combined in a composite fingerprint. Composite signatures combine properties responding to contrasting environmental controls and thereby improve the discrimination of potential sources (Collins & Walling, 2002).

In order to characterize the water quality of the largest rivers of the USA, the US Geological Survey's National Stream Quality Accounting Network (NASQAN) was redesigned in 1994 to estimate the annual mass flux of constituents at a network of fixed stations in the Mississippi, Rio Grande, Colorado, and Columbia river basins. This paper explores the potential for using such routine water quality samples and associated geochemical analyses as a basis for providing a preliminary assessment of the spatial sources of suspended sediment transported by the Ohio River. Since sampling programmes for suspended sediment are resource intensive, opportunities to use the data generated by routine water quality sampling programmes to investigate catchment sediment sources should be explored. In doing so, it is however important to acknowledge that the sampling strategies of routine programmes mean that their water quality data are frequently biased, under-sampling the larger flood events dominating annual sediment transport regimes. Nonetheless, the use of sediment geochemical data from water quality monitoring programmes provides a unique opportunity to undertake preliminary screening of sediment sources in larger river basins. The majority of existing sediment source apportionment work has been undertaken in small (<500 km^2) catchments.

METHODS

The approach

Instantaneous routine suspended sediment samples (Table 1) have been collected as part of the NASQAN programme on the upper Ohio River at Greenup (n = 33), the Wabash River at New Harmony (n = 37), the Cumberland River at Smithland (n = 24) and the Tennessee River at Paducah (n = 45) (Fig. 1). The sets of suspended sediment samples were used to characterise the fingerprint properties of sediment originating from these four tributary sub-catchment spatial sources. Sediment samples (n = 77) collected further downstream on the main stem of the Ohio River near Grand Chain (Fig. 1) were used as the overall outlet samples in the fingerprinting exercise. The entire sets of suspended sediment samples collected for either the individual tributaries or the overall outlet were grouped into respective sampling populations for the purpose of characterising sediment fingerprint properties.

Table 1 Summary of the suspended sediment sampling frequency.

Sampling year	Upper Ohio River at Greenup	Wabash River at New Harmony	Cumberland River at Smithland	Tennessee River at Paducah	Ohio River near Grand Chain
1996	1	1		2	14
1997	11	13		12	14
1998	13	14		11	15
1999	5	4		4	2
2000	1	1		4	5
2001			1	3	4
2002			6	3	4
2003			6	5	4
2004			5		6
2005			3		6
2006			1		3
Start date	04/12/1996	27/11/1996	12/12/2001	21/11/1996	05/02/1996
End date	09/08/2000	13/08/2000	22/06/2006	06/11/2003	03/05/2006

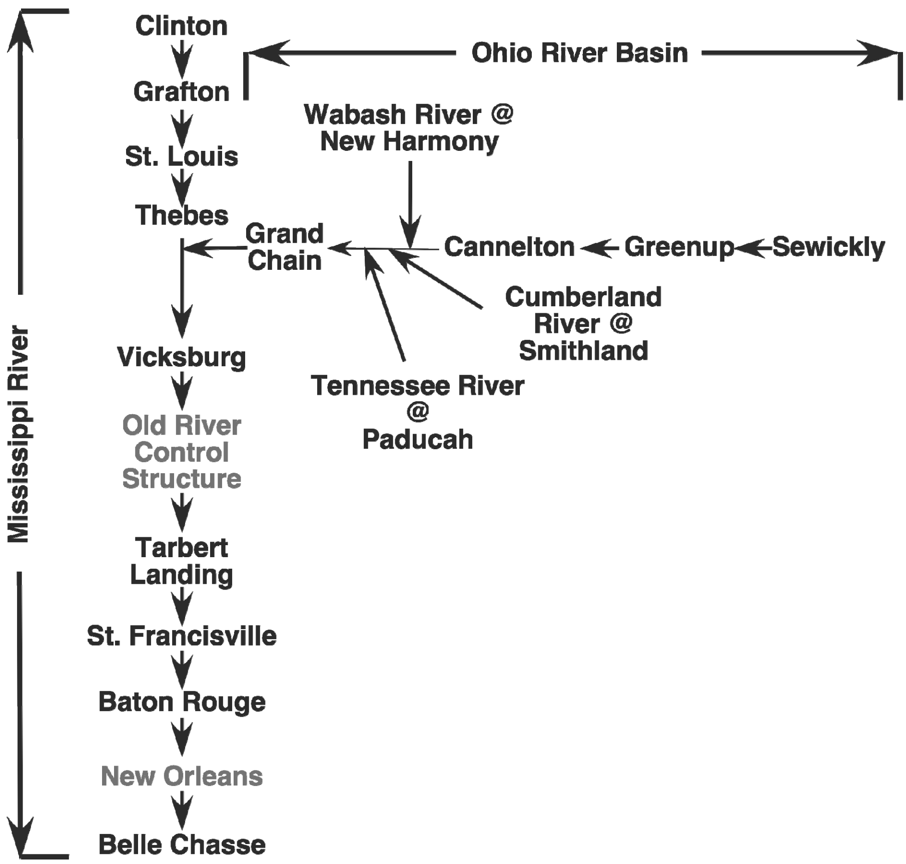

Fig. 1 Schematic of the NASQAN water quality sampling on the Mississippi River.

Sediment sampling and laboratory analyses

In order to provide estimates of annual fluxes of suspended sediment and associated trace elements, between 1.00 to 1.25 g of depth- and width-integrated isokinetic sediment samples were collected at each site throughout the year with an emphasis on wet seasons. Whole-water samples were dewatered and oven dried at 105°C. Sediment recovered from the water samples was analysed for a suite of trace elements (Ag, Cu, Pb, Zn, Cd, Cr, Co, Ni, Ba, V, Li, Be, Mo, Sr, As, Sb, Se, Hg, Fe, Mn, Al and Ti), as well as P, total organic C, total C, total S and total N. Analytical precision typically exceeded ±10%. More detail on the methods used for water sampling and laboratory analyses can be found in Horowitz *et al.* (1989, 2001).

Spatial sediment source characterisation and discrimination

There are clearly some limitations when geochemical data from routine water quality monitoring programmes are used for ascribing the relative contributions of individual sediment sources. Firstly, the sediment properties analysed were not chosen *a priori* for their discriminatory powers for sediment un-mixing, but rather on the basis of the priorities of the overarching water quality monitoring programme. It is therefore possible that the discrimination of sediment sources could be improved with further complementary work. Secondly, sediment sampling locations for characterising the spatial sources are pre-fixed, but, in most instances, the sampling sites will be located in downstream locations on the individual tributaries, thereby providing spatially-integrated sediment samples. Thirdly, the temporal representativeness of the samples must be examined carefully. Instantaneous water samples collected as part of routine monitoring programmes essentially represent snapshots and therefore do not provide time-integrated samples of the sort increasingly used in source fingerprinting investigations (Collins & Walling, 2006; Walling *et al.*, 2008). Time-integrated samples covering periods of weeks to months are temporally more representative of sediment transport and can therefore be treated as discrete samples in source apportionment studies. In situations where instantaneous sediment samples are collected, it is more appropriate to combine all the samples retrieved for specific sources into single sets or sampling populations. Despite such issues, however, the use of sediment geochemical data assembled by routine water quality monitoring programmes permits preliminary investigation of sediment sources in larger river basins without dedicated sampling and analytical costs being incurred.

Of the 27 properties analysed for the NASQAN programme, an optimum composite fingerprint comprising P, Co, Sr, Fe, Cr, Se, Mn and Ti was selected for discriminating the four individual tributary sub-catchment spatial sediment sources on the Ohio River (Table 2). This optimum fingerprint was selected using the minimization of Wilks's lambda as a stepwise selection algorithm within the Discriminant Function Analysis routine of the SPSS software package (Collins *et al.*, 1997). Summary statistics for the properties comprising the optimum composite signature are presented in Table 3. For the selected properties, the coefficients of

Table 2 The optimum composite fingerprint for discriminating tributary sub-catchment spatial sediment sources in the Ohio River basin.

Step	Fingerprint property selected	% tributary sub-catchment suspended sediment samples classified correctly
1	P	68.3
2	Co	77.7
3	Sr	86.3
4	Fe	84.2
5	Cr	90.6
6	Se	93.5
7	Mn	94.2
8	Ti	95.0

Table 3 Summary statistics of selected properties for the source tributaries.

	P mg kg^{-1}	Co mg kg^{-1}	Sr mg kg^{-1}	Fe wt %	Cr mg kg^{-1}	Se mg kg^{-1}	Mn mg kg^{-1}	Ti wt %
Upper Ohio River at Greenup (33*)								
Mean	1177	30	231	4.1	95	1.4	2593	0.40
STD	304	11	133	1.1	26	0.5	973	0.11
Minimum	630	7	110	1.2	48	0.4	860	0.12
Maximum	2100	51	660	5.6	170	2.6	4900	0.59
Wabash River at New Harmony (37*)								
Mean	1249	13	150	3.2	64	1.0	1483	0.37
STD	202	2	54	0.5	10	0.4	582	0.05
Minimum	930	10	100	2.3	46	0.4	860	0.24
Maximum	1900	17	320	4.5	88	2.1	3500	0.45
Cumberland River at Smithland (24*)								
Mean	2158	16	113	3.1	81	0.9	2433	0.46
STD	474	3	26	0.6	16	0.3	585	0.07
Minimum	1200	8	82	1.7	54	0.1	1500	0.27
Maximum	3300	21	200	4.1	120	2.0	3900	0.54
Tennessee River at Paducah (45*)								
Mean	1884	17	157	3.4	128	1.0	3631	0.39
STD	408	3	67	0.7	58	0.3	2249	0.09
Minimum	1200	8	82	1.7	54	0.1	1500	0.27
Maximum	3300	21	200	4.1	120	2.0	3900	0.54

* number of samples used to derive the statistics.

variation (CV) ranged between 26 and 57% for the upper Ohio at Greenup, 12 and 40% for the Wabash River at New Harmony, 15 and 36% for the Cumberland River at Smithland and 20 and 62% for the Tennessee River at Paducah. No systematic trends could be identified in terms of the relative magnitude of elemental variability within each tributary. The CV for the properties at the overall catchment outlet on the Ohio River near Grand Chain was lower (9–24%), possibly reflecting the effect of catchment scale and routing on averaging sediment properties (cf. Klages & Hsieh, 1975; Collins *et al.*, 1998).

Spatial sediment source modelling

Sediment source apportionment modelling is commonly based on the principle of mass conservation, with the following set of constraints (Henry, 1991):

1. values for source composition and source contribution must be non-negative;
2. predicted values for the contribution from any individual source to the catchment outlet sediment samples should be less than 100%;
3. the sum of all predicted source contributions to the catchment outlet sediment samples should equal 100% .

Application of these linear boundary constraints ensures that mixing model solutions are not only mathematically correct, but also physically sound.

A set of linear equations is established to compare the fingerprint properties of the tributary sub-catchment and overall outlet suspended sediment samples. Given the over-determination of the solutions, the relative contributions from the individual spatial sources are estimated by minimizing the sum of squares of the weighted relative errors (Collins *et al.*, 1997):

$$\sum_{i=1}^{n}\left[C_i - \left(\sum_{j=1}^{m} P_j S_{ji} Z_j O_j\right) / C_i\right]^2 \quad (1)$$

where n is the number of fingerprint properties comprising the optimum composite fingerprint, m is the number of tributary sub-catchment spatial sources, C_i is the ith property of suspended sediment samples collected from the overall catchment outlet, S_{ji} is the ith property of suspended sediment collected from tributary sub-catchment spatial source j, and Z and O are source-specific correction factors for particle size and organic carbon content, respectively.

Early source fingerprinting studies used a single measured average value of each property as mixing model input for each source under scrutiny, and thereby did not take explicit account of the uncertainty in estimating source property means based on a small number of samples. Such work yielded sediment source estimates without confidence limits or consideration of the uncertainty therein. More recently, a Monte Carlo approach has been adopted to incorporate the measured variance (uncertainty) associated with the sediment properties comprising composite fingerprints and used to characterise the sources under scrutiny (cf. Collins & Walling, 2007; Wilkinson *et al.*, 2009). A similar, but modified approach was used during this study, incorporating the uncertainty associated with estimating the mean fingerprint properties of both the sources and overall outlet sediment samples using relatively few samples (Collins *et al.*, 2009). Accordingly, the mean and standard deviation (STD) measured for each property for each individual tributary sub-catchment spatial source and the end-member on the basis of the sample sets detailed in Table 1 were used to generate a corresponding population of random numbers, that has a Normal distribution with specified statistics (mean and STD). Random sampling from each population during 5000 repeat mixing model iterations was used to characterise the overall uncertainty range in the predicted mean contributions from the tributary sub-catchments. Previous work has tended to use the frequency distributions of predicted source contributions generated by the repeat iterations to calculate 95% confidence limits around the overall mean. As an alternative to this conventional approach, the results of the Monte Carlo analysis (probability density functions) were used to estimate the weighted average relative inputs (R) from the individual tributary sub-catchment spatial sediment sources:

$$R = \sum_{i=1}^{n} v_i \; F_i \tag{2}$$

where n is the number of intervals for the relative contribution which has a value of 0 to 1; and v and F are the mid-value and the relative frequency for the ith interval, respectively. Use of the weighted average approach afforded a convenient means of summarising the spatial source contributions on the basis of a single number, but taking account of the Monte Carlo analysis. The numerical modelling was undertaken using the Solver function in Microsoft Excel, automated using VBA code.

RESULTS AND DISCUSSION

The relative frequency distributions (probability density functions) for the predicted contributions from each tributary sub-catchment spatial sediment source on the basis of the 5000 mixing model repeat iterations are presented in Fig. 2. These results represent the full uncertainty ranges in the predicted mean contributions from the individual spatial sediment sources and reflect the uncertainty in the fingerprint property means used as input for the numerical mixing model. The estimated source proportions relate to the entire set of suspended sediment samples retrieved from the Ohio River near Grand Chain over the period 1996–2006. Predicted mean relative contributions from the upper Ohio River at Greenup ranged between 0 and 100%, with 41% of the 5000 iterations suggesting an input of ≤10%, and 59% an input of ≤20%. In the case of the Wabash River at New Harmony, the predicted mean relative contributions also ranged between 0 and 100%, with 57% of the repeat solutions predicting an input of ≤30%. For the Cumberland River at Smithland, the predicted mean contributions to sediment samples collected from the main Ohio River near Grand Chain ranged between 0 and 100%, with 57% of the repeat iterations suggesting an input of ≤40%. For the Tennessee River at Paducah, 65% of the repeat realisations suggested a mean relative contribution of ≤20% (Fig. 1). Again, the overall range of the predicted mean relative inputs from this tributary sub-catchment spatial sediment source was 0–100%. On the basis of the probability density functions presented in Fig. 2 and using equation (2), it was estimated that the weighted average relative inputs from the individual tributary sub-catchment spatial sediment sources were in the order: Cumberland River at Smithland (36%) > Wabash River at New Harmony (28%) > upper Ohio River at Greenup (19%) > Tennessee River at Paducah (18%) (Fig. 3).

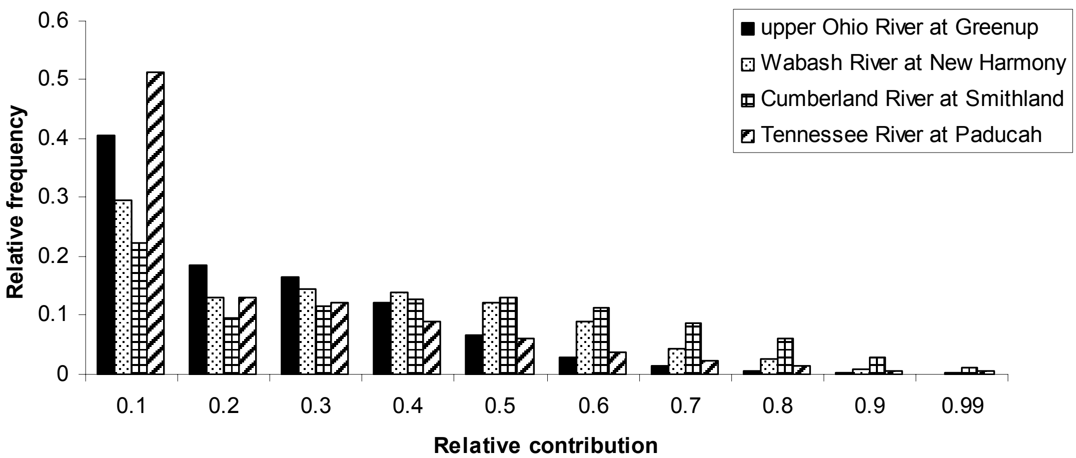

Fig. 2 Frequency distributions of optimised solutions for the mean relative contributions from the individual tributary sub-catchment spatial sediment sources.

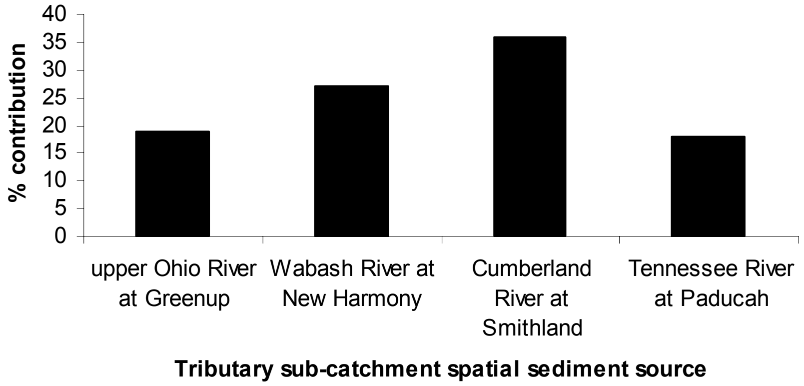

Fig. 3 The weighted average relative contributions (1996–2006) from the individual tributary sub-catchment spatial sediment sources.

Because a Monte Carlo simulation approach was employed, the preliminary results discussed above are, arguably, only one realisation of many other possibilities. The reproducibility of the estimates was therefore assessed using sensitivity tests. Repeat optimisation runs (5000 iterations each) were undertaken using the same source and end-member fingerprint property simulated deviates. These test runs confirmed that the differences between the frequency weighted average relative contributions were less than 1%. Two-sample Kolmogorov-Smirnov tests for differences in the absolute maximum (D_{max}) contrasts further demonstrated that the frequency distributions from repeated optimisation runs each containing 5000 random samples were from the same distribution at the $\alpha = 0.01$ confidence level. It was therefore concluded that the selection of 5000 random deviates for source and end member fingerprint property means was sufficient enough to yield reproducible results for this study.

While sensitivity analysis confirmed the reproducibility of the results using sets of 5000 repeat iterations, it is always desirable to try to minimise the number of random samples during Monte Carlo analysis to improve computing efficiency. Accordingly, the optimisation analysis was repeated using five randomly selected sets of each of 2000, 1000 and 500 iterations. The

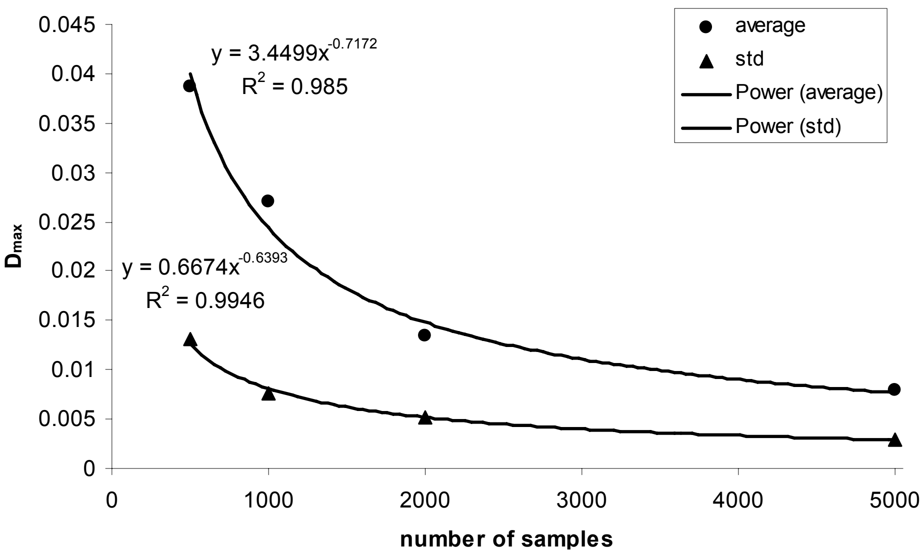

Fig. 4 Relationship between the number of random samples (500, 1000, 2000 or 5000) used during five repeat runs of Monte Carlo analysis and the average and standard deviation of the D_{max} between the predicted source contributions.

results in Fig. 4 suggested that the use of fewer random samples during Monte Carlo analysis tended to increase the average and standard deviation of the maximum difference (D_{max}) between the results of the sets of five repeat optimisation runs, as represented by the corresponding predicted frequency distributions. Therefore, extreme caution should be exercised with respect to reducing the number of random samplings during Monte Carlo analysis since it may lead to instability of the optimisation results.

CONCLUSION

A preliminary assessment of the spatial sources of suspended sediment in the Ohio River basin above Grand Chain was undertaken using NASQAN routine water quality monitoring data (1996–2006) and a Monte Carlo simulation approach. Summary statistics on the fingerprint properties of sediment originating from the spatial sources and passing the overall outlet sampling site were derived using conventional parametric (mean and STD) statistics. The predicted overall mean contributions from the individual spatial sediment sources were estimated using weighted relative frequencies, thereby avoiding the need to cite confidence limits. The interpretation of single values that are still based on the results of uncertainty analysis is likely to be easier for catchment managers. Whilst the limitations of using routine water quality sampling data for sediment sourcing studies should always be borne in mind, such work permits a preliminary screening of sediment provenance in large drainage basins.

Acknowledgements The authors gratefully acknowledge all the field personnel that participated in the NASQAN programme and who collected the requisite large-volume samples used to generate the chemical data used in this study.

REFERENCES

Catt, J. A., Howse, K. R., Farina, R., Brockie, D., Todd, A., Chambers, B. J. Hodgkinson, R., Harris, G. L. & Quinton, J. N. (1998) Phosphorus losses from arable land in England. *Soil Use Manage.* **14**, 168–174.

Collins, A. L. & Walling, D. E. (2002) Selecting fingerprint properties for discriminating potential suspended sediment sources in river basins. *J. Hydrol.* **261**, 218–244.

Collins, A. L. & Walling, D. E. (2004) Documenting catchment suspended sediment sources: problems, approaches and prospects. *Progr. Phys. Geogr.* **28**, 159–196.

Collins, A. L. & Walling, D. E. (2006) Investigation of the remobilization of fine sediment stored on the channel bed of lowland permeable catchments in the UK. In: *Sediment Dynamics and the Hydromorphology of the Fluvial System* (ed. by J. S. Rowan, R. W. Duck & A. Werritty), 471–479. IAHS Publ. 306. IAHS Press, Wallingford, UK.

Collins, A. L., Walling, D. E. & Leeks, G. J. L. (1996) Composite fingerprinting of the spatial source of fluvial suspended sediment: a case study of the Exe and Severn River basins, United Kingdom. *Geomorphologie, Relief, Processus, Environnement* **2**, 41–54.

Collins, A. L., Walling, D. E. & Leeks, G. J. L. (1997) Source type ascription for fluvial suspended sediment based on a quantitative composite fingerprinting technique. *Catena* **29**, 1–27.

Collins, A. L., Walling, D. E. & Leeks, G. J. L. (1998) Use of composite fingerprints to determine the spatial provenance of the contemporary suspended sediment load transported by rivers. *Earth Surf. Proc. Landf.* **23**, 31–52.

Collins, A. L., Walling, D. E., Sichingabula, H. M. & Leeks, G. J. L. (2001) Suspended sediment source fingerprinting in a small tropical catchment and some management implications. *Appl. Geogr.* **21**, 387–412.

Collins, A.L., Walling, D. E., Webb, L. & King, P. (2009). Particulate organic carbon sources and delivery to river channels in the Somerset Levels ECSFDI priority catchment, southwest UK. *Int. J. River Basin Manage.* **7**, 277–291.

Grieg, S. M., Sear, D. A. & Carling, P. A. (2005) The impact of fine sediment accumulation on the survival of incubating salmon progeny: implications for sediment management. *Sci. Total Environ.* **344**, 341–258.

Hatfield, R. G. & Maher, B. A. (2009) Fingerprinting upland sediment sources: particle size-specific magnetic linkages between soils, lake sediments and suspended sediments. *Earth Surf. Processes Landf.* **34**, 1359–1373.

Henry, R. C. (1991) Multivariate receptor models, In: *Receptor Modeling for Air Quality Management* (ed. by P. K. Hopke), 117–147. Elsevier Science Publishers, Amsterdam, The Netherlands.

Horowitz, A. J., Elrick, K. A. & Hooper, R. P. (1989) The prediction of aquatic sediment-associated trace element concentration using selected geochemical factors. *Hydrol. Processes* **3**, 347–364

Horowitz, A. J., Elrick, K. A. & Smith, J. J., (2001) Estimating suspended sediment and trace element fluxes in large river basins: methodological considerations as applied to the NASQAN programme. *Hydrol. Processes* **7**, 1107–1132

Klages, M. G. & Hsieh, Y. P. (1975) Suspended solids carried by the Gallatin River of southwestern Montana: II. Using mineralogy for inferring sources. *J. Environ. Qual.* **4**, 68–73.

Motha, J. A., Wallbrink, P. J., Hairsine, P. B. & Grayson, R. B. (2004) Unsealed roads as suspended sediment sources in an agricultural catchment in south-eastern Australia. *J. Hydrol.* **286**, 1–18.

Packman, A. I. & Mackay, J. S. (2003) Interplay of stream-subsurface exchange clay deposition and stream bed evolution. *Water Resour. Res.* **39**, 4-1–4-9.

Russell, M. A., Walling, D. E. & Hodgkinson, R. A. (2001) Suspended sediment sources in two small lowland agricultural catchments in the UK. *J. Hydrol.* **252**, 1–24.

Wallbrink, P. J., Martin, C. E. & Wilson, C. J. (2003) Quantifying the contributions of sediment, sediment-P and fertilizer-P from forested, cultivated and pasture areas at the landuse and catchment scale using fallout radionuclides and geochemistry. *Soil Till. Res.* **69**, 53–68.

Walling, D. E. (2005) Tracing suspended sediment sources in catchments and river systems. *Sci. Total Environ.* **344**, 159–184.

Walling, D. E., Owens, P. N. & Leeks, G. J. L. (1999) Fingerprinting suspended sediment sources in the catchment of the River Ouse, Yorkshire, UK. *Hydrol. Processes* **13**, 955–975.

Walling, D. E., Collins, A. L. & Stroud, R. (2008) Tracing suspended sediment and particulate phosphorus sources in catchments. *J. Hydrol.* **350**, 274–289.

Warren, N., Allan, I. J., Cater, J. E., House, W. A. & Parker, A. (2003) Pesticides and other micro-organic contaminants in freshwater sedimentary environments – a review. *Appl. Geochem.* **18**, 159–194.

Wilkinson, S. N., Wallbrink, P. J., Hancock, G. J., Blake, W. H., Shakesby, R. A. & Doerr, S. H. (2009) Fallout radionuclide tracers identify a switch in sediment sources and transport-limited sediment yield following wildfire in a eucalypt forest. *Geomorph.* **110**, 140–151.