



Assessing the source and delivery processes of organic carbon within a mixed land use catchment using a combined *n*-alkane and carbon loss modelling approach

C. Wiltshire¹ · M. Glendell² · T. W. Waine¹ · R. C. Grabowski¹ · J. Meersmans^{1,3}

Received: 12 October 2021 / Accepted: 15 March 2022 / Published online: 8 April 2022
© The Author(s) 2022

Abstract

Purpose Understanding fluxes of soil organic carbon (OC) from the terrestrial to aquatic environments is crucial to evaluate their importance within the global carbon cycle. Sediment fingerprinting (SF) is increasingly used to identify land use-specific sources of OC, and, while this approach estimates the relative contribution of different sources to OC load in waterways, the high degree of spatial heterogeneity in many river catchments makes it challenging to precisely align the source apportionment results to the landscape. In this study, we integrate OC SF source apportionment with a carbon loss model (CLM) with the aim of: (i) reducing ambiguity in apportioning OC fluxes when the same land use exists in multiple locations within a catchment; and (ii) identifying factors affecting OC delivery to streams, e.g., buffer zones.

Methods Two main approaches were used in this study: (i) identification of the sources of freshwater bed sediment OC using *n*-alkane biomarkers and a Bayesian-based unmixing model; and (ii) modelling and analysis of spatial data to construct a CLM using a combination of soil OC content modelling, RUSLE soil erosion modelling and a connectivity index. The study was carried out using existing OC and *n*-alkane biomarker data from a mixed land use UK catchment.

Results Sediment fingerprinting revealed that woodland was the dominant source of the OC found in the streambed fine sediment, contributing between 81 and 85% at each streambed site. In contrast, CLM predicted that arable land was likely the dominant source of OC, with negligible inputs from woodland. The areas of the greatest OC loss in the CLM were predicted to be from arable land on steeper slopes surrounding the stream channels. Results suggest extensive riparian woodland disconnected upslope eroded soil OC and, concomitantly, provided an input of woodland-derived OC to the streams. It is likely the woodland contribution to streambed OC is derived from litter and leaves rather than soil erosion.

Conclusion This study demonstrates how location-specific OC sources and delivery processes can be better determined using sediment fingerprinting in combination with CLM, rather than using sediment fingerprinting alone. It highlights that, although wooded riparian buffer strips may reduce the impact of upslope, eroded soil OC on waterways, they could themselves be a source of OC to stream sediments through more direct input (e.g., organic litter or leaf debris). Characterising this direct woodland OC as a separate source within future fingerprinting studies would allow the contributions from any eroded woodland soil OC to be better estimated.

Keywords Terrestrial to aquatic fluxes · Soil organic carbon erosion · Sediment fingerprinting · *n*-alkanes

Responsible editor: J Patrick Lacey

✉ C. Wiltshire
c.wiltshire@cranfield.ac.uk

Extended author information available on the last page of the article

1 Introduction

Soils provide essential ecosystem services, including biomass production, grazing land, forestry, water filtering capacity and, most critically for climate regulation, storage of carbon (Vogel et al. 2018; Wiesmeier et al. 2019). The importance of soil organic carbon (SOC) is widely recognized for soil structure, productivity and the global C cycle.

Soil erosion linked to climate change and human activity threatens the ability of this largely non-renewable resource (Gobin et al. 2004) to continue its vital roles and has detrimental effects on infrastructure and aquatic environments due to excess land to water–sediment fluxes (Bilotta et al. 2008; Rickson 2014; Owens et al. 2016).

The assessment of the relative contribution of different terrestrial sources to organic matter load in waterways can be achieved using sediment fingerprinting (SF) with plant-specific biomarkers (Cooper et al. 2015; Alewell et al. 2016; Zhang et al. 2017; Glendell et al. 2018; Hirave et al. 2020a). Although SF can identify the land use-specific origin of stream OC, it cannot pinpoint the exact origin of that OC if the same land use exists in multiple locations within a catchment—each with different susceptibility to erosion and connectivity to the streams. Source classifications within SF are often too broad (e.g., arable land or forest) to enable precise sources (e.g., specific fields or landscape features) to be determined and management strategies to be targeted (Owens et al. 2016). The ambiguity in OC origin can be reduced if the spatial distribution of erosion prone areas and their likely connection to the streams can be identified. Net catchment erosion can be modelled using sediment delivery models (e.g., WaTEM/SEDEM (Van Oost et al. 2000; Van Rompaey et al. 2001; Verstraeten et al. 2002)); however, accurate predictions for these models require calibration, commonly carried out using outlet sediment yield data (Krasa et al. 2019; Luo et al. 2021). However, sediment yield data are not available for many catchments, and are usually only available at the catchment outlet. For catchments where a lack of sediment yield data may negate the advantage that could be obtained from the application of more sophisticated erosion and routing models, a simple carbon loss model (CLM) can be constructed using spatially distributed carbon sampling (commonly collected for land use-specific SF), together with an empirical erosion model, such as the Revised Universal Soil Loss Equation (RUSLE) (Wischmeier and Smith 1978; Desmet and Govers 1996; Renard et al. 1997), and a connectivity index (CI) which provides an estimate of potential connection between areas of upslope erosion and streams. The extensive literature and data accessibility for RUSLE and CI mean these methods can be easily applied in a wide variety of catchments using available data (Cavalli et al. 2013; ESDAC 2014, 2015a; Panagos et al. 2014, 2015; Alewell et al. 2019).

In Carminowe Creek catchment, Cornwall, UK, there has been an increasing input of woodland-derived organic matter to sediments in Loe Pool (the lake at the catchment outlet) over a period of 60 years (Glendell et al. 2018). Glendell et al. (2018) concluded this increase could be related to enhanced soil erosion, or alternatively, an increase in riparian woodland disconnecting OM inputs from upslope arable land uses. They suggested that coupling fingerprinting with

soil erosion modelling could be a useful tool for quantifying these terrestrial-to-aquatic OC fluxes. To this end, this study quantifies the relative contributions of woodland and arable soil OC to the sediments of multiple sites in streams leading to Loe Pool using the existing *n*-alkane biomarker data of Glendell et al. (2018) with a Bayesian unmixing model SF approach. Sediment fingerprinting estimates were compared with the sources of eroded soil OC reaching the streams estimated using a CLM to assess the origins and delivery processes of streambed OC. The aim of this paper is to evaluate whether location-specific OC sources and delivery processes can be better determined using SF in combination with CLM, rather than SF alone.

2 Material and methods

Two main approaches were used in the study: (i) identification of the sources of freshwater bed sediment OC using *n*-alkane biomarkers and a Bayesian-based unmixing model; and (ii) modelling and analysis of spatial data to construct a CLM using a combination of % SOC content modelling, RUSLE soil erosion modelling and a connectivity index (CI) (Borselli et al. 2008; Cavalli et al. 2013) ($\text{SOC}\% \times \text{RUSLE} \times \text{CI}$).

2.1 Study catchment

Carminowe Creek catchment (4.8 km²) in south-west England, UK (Fig. 1) drains into a large freshwater lake, Loe Pool (0.54 km²). Carminowe Creek comprises two main streams (referred to below as “North” and “South” streams), with a joint outlet to Loe Pool on its eastern side. The mean annual precipitation is ca. 1000 mm and mean annual temperature is ca. 11 °C (Met Office 2021). The catchment bedrock of Devonian mudstone, siltstone and sandstone is overlaid by freely draining loamy soils. The principal land use within the catchment is arable with areas of permanent grassland found on steeper slopes and woodland predominantly found along the riparian corridor (Glendell et al. 2018).

2.2 Samples and analysis

Existing soil and sediment data were used for this research. Full details of sample collection, processing and analysis can be found in Glendell et al. (2018) and are briefly summarised here. Four land uses were considered in their study: arable, temporary grassland referred to as “ley”, permanent grassland referred to as “grassland” and woodland (including riparian woodland areas). Seventy-five soil cores (8 cm diameter, depth 0–15 cm) were taken (30 from arable, 26 from ley, 14 from grassland and five from woodland were

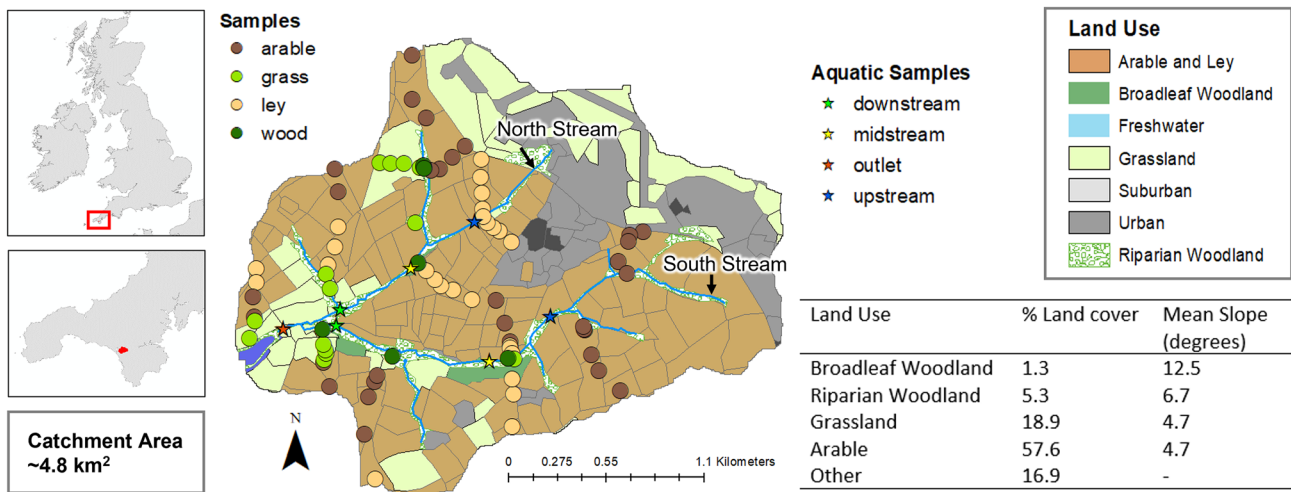


Fig. 1 Carminowe Creek catchment, UK, showing the different land uses and terrestrial and aquatic sample locations (adapted from Glendell et al. (2018)) and a summary of percentage cover and mean slope (in degrees—derived from LiDAR-based digital terrain and surface model for SW Eng-

land [TELLUS SW-Project] ©NERC (Centre for Ecology & Hydrology; British Antarctic Survey; British Geological Survey) (Ferraccioli et al. 2014)) for catchment land uses

collected from 0–15 cm soil depth in summer 2015) (Fig. 1), where required (e.g., in woodland) leaf litter was removed from the surface before sample collection. In addition, streambed sediment samples were collected at six locations along the North and South tributaries (upstream, midstream and downstream) and at the joint catchment outlet. These locations will be referred to as Outlet (OL), North Lower (NL), North Mid (NM), North Upper (NU), South Lower (SL), South Mid (SM) and South Upper (SU). Soil samples were oven dried at 40 °C before sieving. Streambed samples were wet sieved to remove coarse vegetation debris. The criteria for sieving the source and sediment samples were to retain as much soil/sediment as possible while removing anomalously large residual vegetation or sandy/stoney debris. For the streambed sediments, a 250- μ m sieve removed coarser vegetation/sand debris from the finer sediments but as this sieve size would have removed too much soil from the coarser soil samples, these were sieved to 2 mm. In each case, the soil/sediment passing through the sieve was retained for analysis.

All the soil and streambed samples were processed and analysed for total C ($n=75$); however, *n*-alkane concentrations ($\mu\text{g g}^{-1}$ C) were only obtained for a sub-set of soil source samples (eleven from arable, nine from ley, seven from permanent grassland and four from woodland land use) selected on the basis of likely high hydrological connectivity with the streams to characterise for SF (Glendell et al. 2018).

2.3 Software and data maps

The 1 m \times 1 m resolution Digital Elevation Model (DEM) was obtained from LiDAR-based Digital Terrain Model data

for South West England (Ferraccioli et al. 2014). Pit-filling was undertaken in ESRI ArcMap (V10.6) (ESRI 2017) prior to topographic data generation to remove potential processing errors within the DEM. For the estimation of soil loss in RUSLE, the DEM was resampled to a resolution of 20 m \times 20 m. Twenty metres is a typical resolution for DEM in erosion modelling as the processes to be captured by the RUSLE erosion modelling are at a hill-slope scale (Van Oost et al. 2006). The land cover dataset was based on the Centre for Ecology & Hydrology (CEH) Land Cover map of 2015 (LCM2015), as adapted by Glendell et al. (2018) (Fig. 1). Sub-catchments contributing to the seven streambed sediment sample sites were delineated in ESRI ArcMap (V10.6) (ESRI 2017). In addition, a stream buffer representing the land within 20 m of the stream was delineated for each of the sub-catchments (20 m was selected to match the resolution of the CLM maps).

Unless otherwise stated, all statistical analyses were carried out in R (version 3.6.3) (R Core Team 2020) and RStudio (version 1.1.463) (RStudio Team 2018).

2.4 *N*-alkane tracers

Due to their nature (straight-chain hydrocarbons lacking functional groups), *n*-alkanes are stable, long-lived molecules that can survive in the fossil record for millennia (Bush and McInerney 2013) leading to their use as biomarkers in tracing vegetation changes, not only over decades and centuries (Wang et al. 2015; Chen et al. 2017, 2022; Glendell et al. 2018), but in studies of paleoecology and paleoclimatology (Meyers 2003; Glaser and Zech 2005; Zech et al. 2009). *N*-alkanes within sediments are more resistant to degradation

than other organic biomarkers (e.g. sterols, *n*-alkanoic acids, *n*-alkanols). The longer the chain length, the less soluble the *n*-alkane becomes in water, reducing their metabolism by microorganisms (Cranwell 1981; Ranjan et al. 2015). As a result, alkanes of chain-length > C24 are generally resistant to biodegradation (Singh et al. 2012) and are suitable as conservative sediment tracers.

Selection of the sub-set of source soil samples for *n*-alkane analysis in this catchment (Glendell et al. 2018) was originally carried out with the aim of sediment source fingerprinting at the catchment outlet. Therefore, source soil samples within the sub-catchments at the seven streambed sediment sample locations were unevenly distributed. For this reason, all soil samples from the entire catchment were included in the *n*-alkane source apportionment model to characterise the land use sources. To characterise the *n*-alkane distribution within soils under different land use sources, this study used *n*-alkane concentrations ($\mu\text{g g}^{-1}\text{C}$) for chain lengths C15 to C33. *N*-alkanes proxies obtained from the relative abundances of *n*-alkanes were used as “fingerprints”: the relative percentage of *n*-alkanes C27, C29 and C31 (Torres et al. 2014); the C27 to C31 ratio (Puttock et al. 2014); P_{aq} to understand aquatic versus terrestrial plant input (Ficken et al. 2000); the Odd-to-Even Predominance (OEP) (Zech et al. 2013); and the average chain length (ACL) (Fang et al. 2014) were used (Table 1).

2.4.1 Tracer selection

The Bayesian source apportionment model applied in this study, MixSIAR, accounts for the variability in both sources and mixture through uncertain source characterisation and thus offers an advancement on conventional linear models (Stock and Semmens 2016). The geology and soils in this small catchment are uniform, which should minimise within-source group tracer variability due to these factors. The study of Glendell et al. (2018) found that the *n*-alkane tracers could not distinguish well between the arable and ley land uses in this catchment, and therefore, in the tracer selection procedure, these land

uses were combined giving three land use sources arable ($n = 20$), grassland ($n = 7$) and woodland ($n = 4$). *N*-alkane tracers were selected using the following procedure. Firstly, each tracer was assessed for normality using the Kolmogorov–Smirnov test. A “range test” was then carried out by comparing boxplots of each potential tracer in source samples against the mixtures (streambed sediment) to assess if the range of values for the streambed sediments was within the full range of values for the terrestrial land use sources. The boxplots were created in Excel with the “full range” defined by the whiskers (extending up from the top of the box to the largest data element that is ≤ 1.5 times the interquartile range (IQR) and down from the bottom of the box to the smallest data element that is > 1.5 times IQR); values outside this range were considered outliers. The full range (excluding outliers) was used to account for the small sample sizes available to characterise each land use and the variability in the source fingerprints. The streambed sediment mixtures are represented by a single measurement in each case without any knowledge of the potential mean and distribution. It is therefore possible that the single measurement represents either a value close to the maximum or minimum of the possible tracer values rather than the mean and therefore selecting tracers based on the means and inter-quartile range of the sources was considered too restrictive. Finally, a Kruskal–Wallis non-parametric test followed by a post hoc Dunn test was employed to determine if the tracers could distinguish between the three land use sources.

2.4.2 Virtual mixtures

Once a suitable set of *n*-alkane tracers had been selected, land use discrimination was assessed using a “virtual” mixture with 50/50 contributions from each source by taking the mean of the two sources to represent a 50% contribution from each. Errors were calculated as mean absolute differences between the modelled (MixSIAR) and virtual mixture composition.

Table 1 Description of *n*-alkane derived ratios considered tracers within the MixSIAR sediment fingerprinting source apportionment model for Carminowe Creek catchment, UK

<i>n</i> -alkane ratios	Description
$\%C_{27}, \%C_{29}, \%C_{31}$	Percentage of alkane “i” where: $\%C_i = \frac{C_i}{(C_{27}+C_{29}+C_{31})}$ (Torres et al. 2014)
C_{27}/C_{31}	C27 to C31 ratio estimating the proportion of wood to grass derived organic matter (Puttock et al. 2014)
$PAQ = \frac{C_{23}+C_{25}}{C_{23}+C_{25}+C_{29}+C_{31}}$	P_{aq} to understand aquatic versus terrestrial plant input (Ficken et al. 2000)
$OEP = \frac{C_{27}+C_{29}+C_{31}+C_{33}}{C_{26}+C_{28}+C_{30}+C_{32}}$	Odd-over-even predominance (OEP). (Zech et al. 2013)
$ACL = \frac{27 \times C_{27} + 29 \times C_{29} + 31 \times C_{31} + 33 \times C_{33}}{C_{27} + C_{29} + C_{31} + C_{33}}$	Average Chain length (Fang et al. 2014)

2.5 Bayesian unmixing model (MixSIAR) implementation

MixSIAR uses the mean and standard deviation to characterise tracer properties. MixSIAR is “fully Bayesian” (source data fit hierarchically) and estimates the ‘true’ source means and variances used in the mixture likelihood. Source means and standard deviations used in the mixture likelihood are allowed to deviate from the user-specified values with the amount of deviation dependent on source sample sizes. Estimates of sediment proportions are made using Markov chain Monte Carlo (MCMC) simulations. A full description of this model can be found in Stock and Semmens (2016) and Stock et al. (2018). MixSIAR was formulated using a residual error term and an uninformative prior in all model runs. The MCMC parameters were set as follows: chain length = 100,000, burn = 50,000, thin = 50, chains = 3 (convergence of all models was evaluated using the Gelman-Rubin diagnostic).

2.6 Carbon loss model

A carbon loss model (CLM) was constructed as follows:

$$CLM = SOC\% \times SL \times CI \quad (1)$$

where SOC% is a map of the soil organic carbon content (%), *SL* is a soil loss map constructed using RUSLE and *CI* is a map of connectivity index as defined by Borselli et al. (2008) and Cavalli et al. (2013).

2.6.1 SOC content mapping

To map soil OC (%) across the catchment, soil samples were interpolated using a linear regression model implemented in R (version 3.6.3) (R Core Team 2020) packages “raster” (Hijmans 2020), “sp” (Pebesma and Bivand 2005) and “gstat” (Pebesma 2004). Seven land use and topographic environmental predictor maps were generated using ESRI ArcMap (V10.6) (ESRI 2017): land use, slope, curvature, flow length (longest upslope distance along the flow path, from each cell to the top of the drainage divide), accumulated flow (accumulated weight of all cells flowing into each downslope cell), topographic wetness index (Mayer et al. 2019) and aspect (i.e. compass direction that the steepest slope is facing at a given location). The land uses considered within the model were grassland, arable (a combination of arable and temporary grassland or ley), broadleaf woodland and riparian woodland as these were the land uses available on the land use map adapted from Glendell et al. (2018). Climate and soil parameters were not considered predictors, as, except for one sample, all samples were taken on the same

soil type and climate data were not expected to vary significantly across this small catchment (< 5 km²). The model was selected by highest adjusted *R*² and lowest Akaike Information Criterion (AIC) (Meersmans et al. 2012). A leave-one-out cross-validation was implemented, and the root mean square error (*RMSE*) and *R*² of the model simulations were calculated to check model accuracy against observations.

2.6.2 Connectivity index

To define the degree of connectivity between upslope sediment sources and catchment streams, CI was calculated using the method of Cavalli et al. (2013) and the catchment DEM using ESRI ArcMap (V10.6) (ESRI 2017). For use as a weighting with the soil organic carbon content and RUSLE, CI was re-scaled from 0 to 1.

2.6.3 Soil loss modelling

Long-term average annual soil loss in RUSLE is calculated as:

$$SL = R.K.L.S.C.P \quad (2)$$

where *SL* is the mean soil loss (t ha⁻¹ year⁻¹), *R* is the rainfall intensity factor (MJ mm ha⁻¹ h⁻¹ year⁻¹), *K* is the soil erodibility factor (t ha h ha⁻¹ MJ⁻¹ mm⁻¹), *S* and *L* are the slope and slope-length factors, *C* and *P* are the dimensionless cover-management factor and conservation support practice factor that are heavily dependent on the land use and management. In this small catchment (< 5 km²), single values were used for the RUSLE *R* and *K* factors, based on existing derived spatial datasets (*R* ESDAC 2015a, b; Panagos et al. 2015) (*K* ESDAC 2014; Panagos et al. 2014) (Table 2). A *C* factor map was created by assigning literature values for arable land, grassland, forest and urban areas to the land cover map (Sect. 2.4) (Van Rompaey and Govers

Table 2 RUSLE factors used to estimate long-term average annual soil loss for Carminowe Creek catchment, UK. *R* rainfall intensity factor (MJ mm ha⁻¹ h⁻¹ year⁻¹), *K* soil erodibility factor (t ha h ha⁻¹ MJ⁻¹ mm⁻¹), *S* and *L* slope and slope-length factors, *C* and *P* dimensionless cover-management factor and conservation support practice factor

RUSLE factor	Value
<i>R</i>	586.85 (MJ mm ha ⁻¹ h ⁻¹ year ⁻¹)
<i>C</i> _{arable}	0.21 (Range 0.12–0.34)
<i>C</i> _{forest}	0.005 (Range 0.01–0.001)
<i>C</i> _{grassland}	0.0625 (Range 0.2–0.005)
<i>C</i> _{urban}	0.005 (Range 0.01–0.001)
<i>K</i>	0.041 (t ha h ha ⁻¹ MJ ⁻¹ mm ⁻¹)
<i>P</i>	1

2002; Bakker et al. 2008; Gadiga and Martins 2015; Oliveira et al. 2015) (Table 2). The conservation support practice factor (P) was not considered in this study and was set to 1. The RUSLE LS factor was generated from the DEM in R (version 3.6.3) (R Core Team 2020) using packages “raster” (Hijmans 2020) and “RSAGA” (Brenning et al. 2018: version 7.6.3, method “Desmet and Govers”). The RUSLE factor maps and %SOC map were used to calculate SOC loss using packages “raster”, “RSAGA” and “rgdal” (Bivand et al. 2019).

2.6.4 Land use-specific distribution of carbon loss

The value of C-factor within RUSLE model can be used to account for the differences in erosion potential between land uses. However, the range of values found for the C factor in the literature (Table 2) can lead to a one or two orders of magnitude difference in RUSLE output. Therefore, it was important to evaluate the magnitude of the errors associated with the RUSLE C-factor as well as that introduced by the modelling of SOC content (%SOC) using a Monte Carlo analysis with 3000 iterations. The RUSLE C factor was sampled from a uniform distribution defined by the maximum and minimum values found in the literature (Table 2) and %SOC content was sampled from a uniform distribution defined by ± 1 RMSE from the leave-one-out cross-validation of the %SOC content map (see Sect. 2.6.1). At each iteration the SOC loss from arable and woodland, land uses were calculated, generating a probability distribution for comparison with sediment source proportions estimated using SF.

3 Results

3.1 *N*-alkane distribution

The *n*-alkane distribution of the samples in this catchment is discussed in Glendell et al. (2018) and is summarised here. As expected, C27 and C29 dominated the woodland *n*-alkane distribution (63%) with a smaller contribution from those homologues’ dominant in grasslands (C31, C33—combined proportion of 28%) (Fig. 2a). Conversely, the arable land use *n*-alkane distribution was dominated by contributions from C31, C33 homologues’ dominant in grasslands (combined proportion of 56%) with smaller contributions from those homologues’ dominant in woodlands (C27, C29—combined proportion of 36%). Both land uses had a much smaller contribution from homologues’ dominant in lower plants and mosses (C23, C25—combined proportion of 8–9%). The relative proportions of the *n*-alkane homologues in the streambed sediments were dominated by C27 and C29 (combined proportion of 67–71% for all streambed sites except OL which had a slightly lower proportion of 62%) (Fig. 2a).

3.2 Source apportionment

To evaluate whether location-specific OC sources and delivery processes can be better determined using SF in combination with a CLM, rather than SF alone, the relative contributions of woodland and other land uses were first quantified using the *n*-alkane biomarker data and MixSIAR.

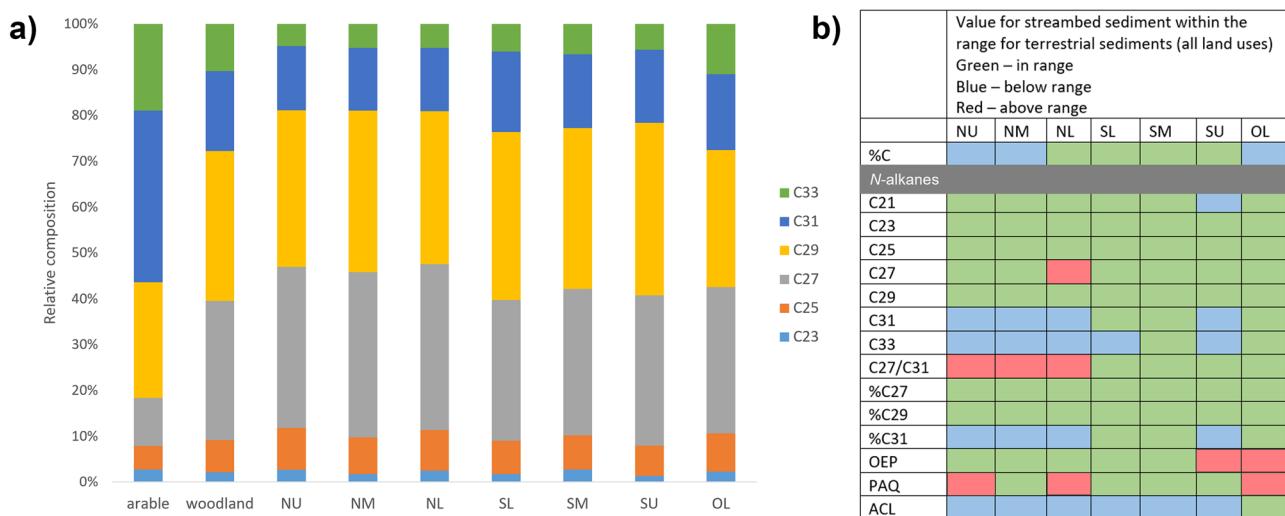


Fig. 2 a Relative mean concentration (%) for mid and long-chain *n*-alkane homologues for the soils of land uses, arable and woodland and streambed sediments OL, NU, NM, NL, SU, SM and SL, and b Range comparison for %C, mid and long-chain *n*-alkane homologues,

and *n*-alkane ratios between terrestrial land uses and streambed sediments OL, NU, NM, NL, SU, SM and SL for the Carminowe Creek catchment, UK

The Kolmogorov–Smirnov test revealed that all *n*-alkane tracers (C27/C31 ratio, %C27, %C29, %C31, OEP, PAQ and ACL) were not significantly different from a normal distribution ($p > 0.05$). The Kruskal–Wallis non-parametric test for the three land use sources (arable, grassland and woodland) revealed that the distribution of *n*-alkane tracers was not the same for every land use ($p < 0.05$) for all tracers except OEP. However, the post hoc Dunn test which compares the land uses pairwise revealed the *n*-alkane tracers could not distinguish grassland from arable. Consequently, as this study is essentially concerned with the relative contribution of woodland and “non-woodland” sources, the grassland and arable data were combined into one source which will henceforth be referred to as “arable”. Analysing this combined arable source, the Kolmogorov–Smirnov test revealed that all *n*-alkane tracers were not significantly different from a normal distribution ($p > 0.05$), except for C27/C31 ratio ($p = 0.022$). The Kruskal–Wallis non-parametric test for the two land use sources (arable and woodland) revealed that the distribution of *n*-alkane tracers was not the same for every land use ($p < 0.05$) for all tracers except OEP. OEP was therefore excluded as a tracer. The range test revealed that for %C27 and %C29, the range of values for the streambed sediments was within the full range of values for the terrestrial land use sources and these two

tracers were therefore selected for use in source apportionment (Fig. 3). The difference in range between the streambed sediment *n*-alkanes and those of the terrestrial land uses was primarily due to the relative depletion of %C31 in the streambed sediments (Fig. 2b) which commensurately reduced the average chain length (ACL) and increased the C27/C31 ratio. The values of the *n*-alkane proxy for aquatic versus terrestrial plant input (PAQ) were generally within the range of the woodland (PAQ 0.12–0.17); however, a few sample sites had larger PAQ values (0.19–0.2). Ankit et al. (2022) ascribe PAQ values < 0.1 to terrestrial vegetation and 0.1–0.4 to emergent macrophytes which could suggest some input of *n*-alkanes from the latter in streambed sediments. However, the woodland PAQ values are also all above 0.1 and it unlikely that emergent macrophytes would make a significant contribution to terrestrial soils. Using MixSIAR and the selected *n*-alkane tracers (%C27 and %C29) land use discrimination was assessed using a “virtual” mixture. The mean absolute difference between the modelled (MixSIAR) and virtual mixture composition was only 0.2% suggesting *n*-alkane tracers %C27 and %C29 could discriminate well between the two land use sources.

Source apportionment using MixSIAR with *n*-alkane tracers %C27 and %C29 found the dominant OC source at every streambed site was woodland. There was

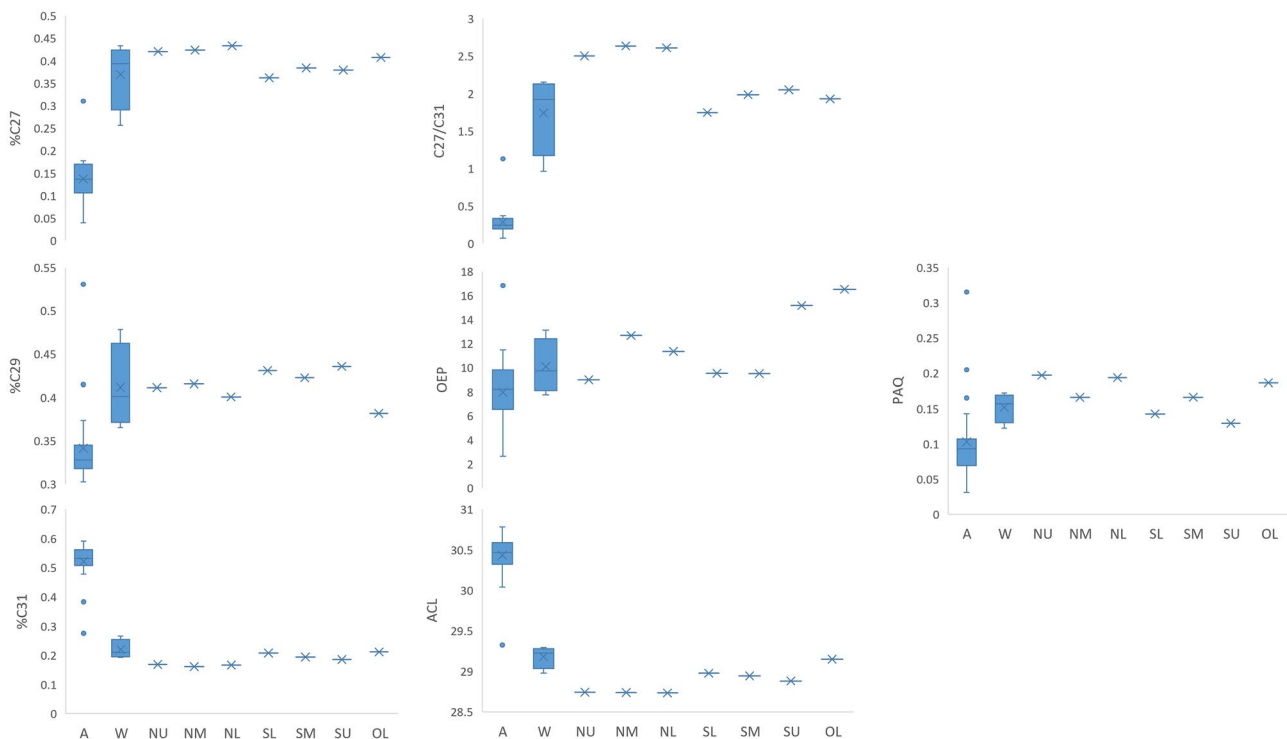


Fig. 3 Box plots of *n*-alkane ratios for the soils of land use types, arable (A), and woodland (W) and streambed sediments OL, NU, NM, NL, SU, SM and SL for the Carminowe Creek catchment. The middle line of the box represents the median and the “x” the mean. Where

present, the box represents the first to third quartile and the whiskers extend from minimum to maximum values excluding outliers (blue dots)

Table 3 Proportion of woodland soil OC input to the streambed sediments OL, NU, NM, NL, SU, SM and SL for the Carminowe Creek catchment estimated using *SF* sediment fingerprinting and a *CLM* carbon loss model at a sub-catchment and 20-m stream buffer scale

	Estimated proportion of woodland soil OC (% ± 1SD)						
	OL	NL	NM	NU	SL	SM	SU
SF	81.64 ± 13.14	85.07 ± 11.58	85.47 ± 11.03	84.61 ± 11.70	81.39 ± 12.68	82.23 ± 12.43	83.29 ± 11.79
CLM (sub-catchment)	0.81 ± 0.52	0.40 ± 0.26	0.34 ± 0.22	0.65 ± 0.42	1.37 ± 0.88	1.17 ± 0.77	0.38 ± 0.26
CLM (20 m buffer)	3.64 ± 2.24	1.39 ± 0.89	0.98 ± 0.65	1.21 ± 0.76	7.68 ± 4.43	3.18 ± 2.04	1.11 ± 0.74

little difference between the seven streambed sites with woodland contributing between 81 and 85% at each site (Table 3).

3.3 Carbon loss modelling

The %OC of the samples in this catchment are discussed in Glendell et al. (2018) and are summarised here. The mean %OC was the greatest within woodland ($7.80 \pm 1.98\%$) land use followed by grassland ($5.40 \pm 1\%$), ley ($3.77 \pm 1.01\%$) and arable land use ($3.05 \pm 0.61\%$). In general, the %OC content of streambed sediments was lower than that of terrestrial land uses, with the highest %OC in streambed sediments (Site SM 3.7%) comparable to that of ley and arable soils. The lowest %OC were found at sites NU (1.16%) and OL (1.55%), which had relatively little woodland nearby, with NU being surrounded by arable and grassland and OL located near steeply sloping grasslands. The largest %OC was seen at site SM (3.68%), which is located next to an extended area of woodland.

The CLM required the spatial distribution of soil OC and to this end the %OC across the catchment was predicted by interpolating %OC of each soil sample using linear regression (Table 4). Land use showed the strongest significant relationship ($p < 0.05$) with %OC (adjusted $R^2 = 0.54$). OC content showed weak significant relationships ($p < 0.05$) with curvature (adjusted $R^2 = 0.07$), TWI (adjusted $R^2 = 0.12$), flow length (adjusted $R^2 = 0.18$) and accumulated flow (adjusted $R^2 = 0.13$); however, when considered together with land use, none of these other covariates was significant. No significant relationships with %OC were found for the other covariates (slope and

aspect). The highest adjusted R^2 (0.54) and lowest AIC were obtained when land use alone was used as a predictor (Table 4). The leave-one-out cross-validation checking model accuracy against observations had a root mean square error (*RMSE*) of 1.35 and R^2 of 0.43. The land uses considered within the model were grassland, arable (a combination of arable and ley), broadleaf woodland and riparian woodland (Sect. 2.3). The highest SOC content was predicted in broadleaved woodland (7.29%), followed by grassland (5.76%), riparian woodland (5.26%) and arable land (3.17%).

The combined CLM ($\text{SOC}\% \times \text{RUSLE} \times \text{CI}$) reveals areas of the greatest OC loss are predicted in arable land on the relatively steeper slopes surrounding the stream channels (Fig. 4). In each of the seven sub-catchments of Carminowe Creek (OL, NU, NM, NL, SU, SM and SL), the proportion of woodland soil OC input to the streambed sediments was estimated using the CLM at a sub-catchment and 20-m stream buffer scale (Table 3). The two scales (sub-catchment and 20-m stream buffer) were used to assess if streambed OC proportions were more aligned with local riparian conditions, rather than those in the wider sub-catchment. At the sub-catchment scale, woodland represents only 6% to 9% of the total land use for each streambed sediment site. This percentage rises at the 20 m buffer scale (37–58%) as most of the woodland is located in close proximity to the streams. The CLM estimated that woodland soil OC represented a relatively small proportion of eroded soil OC likely to reach the streams (< 1.4 at a sub-catchment scale and up to up to $7.7 \pm 4.4\%$ at a 20-m stream buffer scale) with the overwhelming majority originating in arable land.

Table 4 SOC content regression relationship and root mean square error (*RMSE*) and R^2 value resulting from leave-one-out cross-validation. In the context of the linear regression relationship, the variables “grass-

land”, “riparian” and “woodland” are dummy variables which are equal to one when that land use is present and zero otherwise

SOC content regression relationship	<i>RMSE</i>	R^2
$\text{SOC}\% = 3.1694 + 2.5950(\text{grassland}) + 2.0931(\text{riparian}) + 4.1166(\text{woodland})$	1.35	0.43

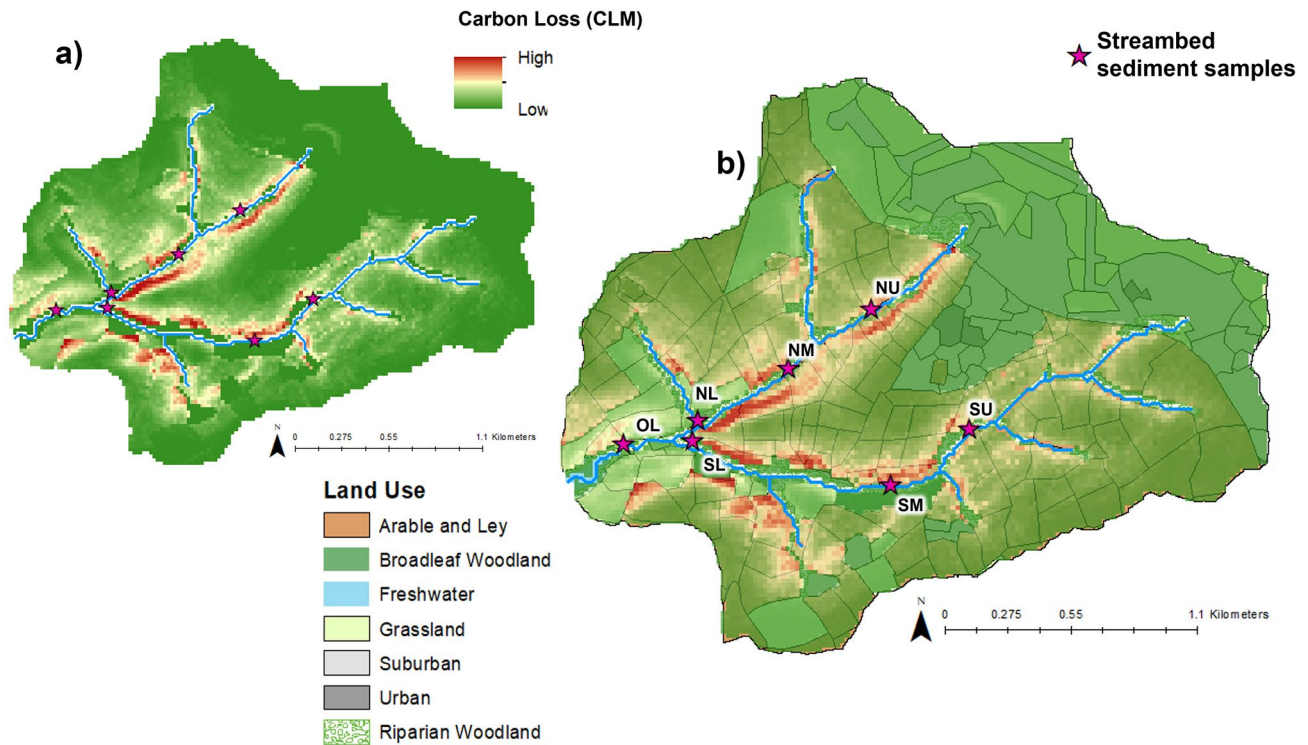


Fig. 4 a Carbon loss model (CLM) and b combined CLM and land use map for Carminowe Creek catchment, UK

4 Discussion

We combined a CLM with SF to characterise OC distribution in soils under different land uses and to quantify the sources of OC in Carminowe Creek, a small, mixed land use, UK catchment.

The CLM predicted areas of the greatest OC loss in arable land on the relatively steeper slopes surrounding the stream channels. The proportion of woodland soil OC input to the streambed OC estimated by CLM at a sub-catchment scale, < 1.4% is smaller than would be expected given its area coverage (6–9%), close proximity to the streams (high connectivity), and relatively high %OC (5.26–7.29 cf. 3.17–5.76% for arable). This is due mainly to a greater protection from erosion afforded by the permanent vegetation found in woodland compared to arable land which has more variable vegetation cover due to human-induced processes (Poesen 2018). This is reflected in the RUSLE C-factor which is much higher (resulting in a significantly higher level of estimated erosion) for arable land than woodland (arable 0.12–0.34, woodland 0.01–0.001). In addition, some of the arable land in this catchment is located on the steep slopes leading down to the stream network which is likely to increase both the speed, and the erosive potential of water runoff and increase the probability of eroded sediment reaching the streams (Renard et al. 1997). The proportion of woodland soil OC input to the streambed OC estimated by

CLM at a 20-m stream buffer scale is larger (up to 7.7%) due to the larger proportion of woodland at this scale (37–58%) but is still significantly smaller than the contribution from arable land due to the higher levels of erosion predicted for that land use. There is a large discrepancy between the CLM estimates of woodland soil OC contributions to streambed OC and those estimated by SF. Neither the carbon loss estimated in close proximity to the streams (CLM 20-m stream buffer scale), nor that in the wider catchment, came close to the 81–85% woodland contribution estimated by SF. The discrepancy between the CLM estimates of woodland soil OC contributions to streambed OC and those estimated by SF suggests that woodland soil is being input to streams by processes not modelled by the CLM, and/or there is a source of woodland vegetation biomarkers not originating from soil.

Carminowe Creek has extensive riparian woodland. This riparian woodland vegetation can reduce delivery of upslope fine-grained sediment to streams (Grabowski and Gurnell 2016; Wu et al. 2021) and, therefore, Carminowe Creek's extensive riparian woodland is likely to have reduced the presence of eroded arable soil OC in the creek bed sediments. As SF estimates OC source contributions directly from streambed sediments, it represents a combination of both potential contribution from upslope terrestrial sources and processes within the stream channel and riparian zone. Terrestrial to aquatic fluxes of OC can originate in this

active and dynamic river “corridor”, which encompasses both the active stream channel and the riparian zone (Wohl et al. 2017) through direct input (e.g. organic litter or leaf debris) and overflow of river banks and the riparian zone (Bright et al. 2020). Bank erosion could, therefore, have contributed woodland soil to the streams. However, a recent assessment of branched tetraether lipids (membrane lipids of soil bacteria) in Carminowe Creek suggested the absence of a clearly recognizable soil brGDGT (branched glycerol dialkyl glycerol tetraethers) signal in creek bed sediments could be explained if there was a relatively limited input of soil material into the creek (Guo et al. 2020). Lewis et al. (2021) found the amount of wood in streams was best explained by riparian tree canopy cover and the length of tree-lined channel upstream. There could be leaves/needles directly associated with this deposited woody debris and its presence in the stream channel can capture additional leaf litter and/or twigs (Lewis et al. 2021).

Hirave et al. (2020b) found little or no difference between *n*-alkane concentrations between fresh plant biomass and the soil organic horizon (O horizon) suggesting that it may be difficult to distinguish between *n*-alkane signatures from those two sources. Stout (2020) found the average chain length and OEP (odd–even predominance) of fresh mature leaves increased and decreased respectively in the corresponding leaf litter and further in the corresponding soil, which they attributed to preferential and progressive degradation of the more abundant C27/C29 homologues relative to the less abundant C31/C33. As a result, OEP is relatively higher and %C31 relatively lower for leaves/litter compared to the more degraded OM in the associated soil. In this study, when comparing the streambed sediment to the terrestrial soils, the OEP values of streambed sediments were similar to or greater, and the %C31 values similar or lower. Direct input of leaf/wood organic matter to the stream sediments could explain the respectively higher and lower OEP and %C31 values of these sediments. Characterising this direct woodland OC as a separate source within future fingerprinting studies would allow the relative contributions from this more direct source and any eroded woodland soil OC to be estimated. This may require the inclusion of biomarkers of plant, fungal and bacterial origin to provide a fingerprint more characteristic of the soil rather than just the vegetation. Although the bacterial brGDGT biomarkers of Guo et al. (2020) were not found to be land use-specific, other biomarkers, such as fatty acids considered common to prokaryotic and eukaryotic organisms, have been found relevant for land use discrimination (Ferrari et al. 2015).

Monte Carlo techniques were used to propagate uncertainties in both the SF and CLM estimates of land use contributions to streambed OC. As the study is concerned with relative contributions from land use sources, the CLM uncertainty analysis was concentrated on factors that were

strongly land use dependent (RUSLE C-factor and OC spatial modelling). The uncertainties associated with the other RUSLE factors and CI were considered independent of land use. Uncertainties in SF results can arise due to factors affecting source and sediment characterisation such as sample size and particle size fractions. The sample size for characterising streambed sediment and woodland in this study was small (only one sample for each streambed site and four samples for the woodland soil). The authors recommend as large a sample size as possible (within practicality and budget constraints) to facilitate a more robust characterisation of the distributions of both soil sources and streambed mixtures resulting in a more robust range test of tracer conservativeness. Finer, lighter particles are more likely to be mobilised by water in the terrestrial environment, and therefore, as in this study, aqueous sediments may end up with a finer particle size distribution than terrestrial sediments. The particle size fractions of the soil and sediment samples used in this study were determined by using sieve sizes that retained as much soil/sediment as possible, while removing anomalously large debris. This resulted in different size fractions, < 2 mm and < 250 µm respectively, for the soil and (relatively finer) streambed sediments. In the study of Geng et al. (2019), the distribution and preservation of *n*-alkanes was found to differ between coarse (> 250 µm) particulate organic matter (POM) and fine POM (< 250 µm). The coarse POM had a greater abundance of plant-derived *n*-alkanes ($n > 20$) with chain-length shortening in the fine POM fraction suggesting a stronger decomposition of *n*-alkanes in that fraction. The respectively higher and lower OEP and %C31 values found for the Carminowe Creek streambed sediments (< 250 µm) indicate less degradation than the coarser (< 2 mm) soil sediments and, therefore, there is unlikely to be an effect due to particle size similar to that found by Geng et al. (2019). It is generally accepted that OM (including *n*-alkanes) are preferentially associated with the finer particle size fractions (< 63 µm) (Quenea et al. 2004; Quéneá et al. 2006). In addition, studies have found the majority of OC resides in the finer soil fractions (Yu et al. 2019; De Mastro et al. 2020). This finer fraction was present in both soil and sediment samples, however, runoff from eroding landscapes can be enriched in these finer, clay sized particles (Starr et al. 2000; Nitzsche et al. 2022) and could have affected the *n*-alkane distribution of streambed sediments relative to the source samples in this study (Lacey et al. 2017). In future studies, analysing terrestrial source soils at different particle size fractions could help quantify any effects on *n*-alkane distributions due to this factor. Uncertainty could be further reduced by using different methods to isolate the finer fraction within the soil samples. Under field conditions, various mechanisms cause soil aggregates to break apart creating finer particle fractions; disintegration of aggregates is a complicated mixture of mechanical

(raindrop impact, field traffic/tillage, roots, earthworms) and hydraulic stresses (Felde et al. 2021). Therefore, using different methods to isolate the finer fraction within the soil samples could highlight any differences in biomarker distribution due to breaking down aggregates using methods such as dry crushing (along more “natural planes of mechanical weakness” i.e., those likely to fail in the field (Felde et al. 2021)) compared to wet/dry sieving and/or sample grinding.

5 Conclusions

This study revealed that combining a CLM with SF enhanced the understanding of the fate of eroded OC and terrestrial to aquatic fluxes for a mixed land use catchment. The results of this study support others that found riparian buffers reduced soil OC loss from terrestrial to aquatic ecosystems (Zhang et al. 2010; Valera et al. 2019; Liu et al. 2020). The approach has highlighted that the amount of upslope OC erosion cannot be reliably equated with delivery to streams unless (i) sites of intermediate storage or “buffers” are also considered (Trimble 1983; Owens 2020), and (ii) estimates of other plant-derived OC sources e.g., direct input of leaf/wood organic matter can be made. It is likely that, although wooded riparian buffer strips may reduce the impact of upslope, eroded soil OC on waterways, they could themselves be a source of OC to stream sediments through more direct input (e.g., organic litter or leaf debris). Characterising this direct woodland OC as a separate source within future fingerprinting studies would allow the contributions from any eroded woodland soil OC to be better estimated. This study was focused on streambed sediments and therefore, average, longer-term OC fluxes. In future studies, it will be important to assess suspended sediment as well as bed sediments to assess any seasonal changes in terrestrial OC origins and delivery processes.

Funding This work was supported by the Natural Environment Research Council and the Biotechnology and Biological Sciences Research Council (Grant number NE/R010218/1) through a studentship award to CW by STARS (Soils Training And Research Studentships) Centre for Doctoral Training and Research Programme; a consortium consisting of Bangor University, British Geological Survey, Centre for Ecology and Hydrology, Cranfield University, James Hutton Institute, Lancaster University, Rothamsted Research and the University of Nottingham, and the CASE funder Scottish Environment Protection Agency (SEPA).

Data availability This study was carried out using existing datasets available from the authors of (Glendell et al. 2018).

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alewell C, Birkholz A, Meusburger K et al (2016) Quantitative sediment source attribution with compound-specific isotope analysis in a C3 plant-dominated catchment (central Switzerland). *Biogeosciences* 13:1587–1596. <https://doi.org/10.5194/bg-13-1587-2016>
- Alewell C, Borrelli P, Meusburger K, Panagos P (2019) Using the USLE: chances, challenges and limitations of soil erosion modelling. *Int Soil Water Conserv Res* 7:203–225. <https://doi.org/10.1016/j.iswcr.2019.05.004>
- Ankit Y, Muneer W, Gaye B et al (2022) Apportioning sedimentary organic matter sources and its degradation state- inferences based on aliphatic hydrocarbons amino acids and $\delta^{15}N$. *Environ Res* 205. <https://doi.org/10.1016/j.envres.2021.112409>
- Bakker MM, Govers G, van Doorn A et al (2008) The response of soil erosion and sediment export to land-use change in four areas of Europe: the importance of landscape pattern. *Geomorphology* 98:213–226. <https://doi.org/10.1016/j.geomorph.2006.12.027>
- Bilotta GS, Brazier RE, Haygarth PM et al (2008) Rethinking the contribution of drained and undrained grasslands to sediment-related water quality problems. *J Environ Qual* 37:906–914. <https://doi.org/10.2134/jeq2007.0457>
- Bivand R, Keitt T, Rowlingson B (2019) rgdal: bindings for the “geospatial” data abstraction library. R package.
- Borselli L, Cassi P, Torri D (2008) Prolegomena to sediment and flow connectivity in the landscape: A GIS and field numerical assessment. *CATENA* 75:268–277. <https://doi.org/10.1016/j.catena.2008.07.006>
- Brenning A, Becker M, Bangs D (2018) RSAGA: SAGA geoprocessing and terrain analysis. R package
- Bright CE, Mager SM, Horton SL (2020) Catchment-scale influences on riverine organic matter in southern New Zealand. *Geomorphology* 353:107010. <https://doi.org/10.1016/j.geomorph.2019.107010>
- Bush RT, McInerney FA (2013) Leaf wax n-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim Cosmochim Acta* 117:161–179. <https://doi.org/10.1016/j.gca.2013.04.016>
- Cavalli M, Trevisani S, Comiti F, Marchi L (2013) Geomorphometric assessment of spatial sediment connectivity in small Alpine catchments. *Geomorphology* 188:31–41. <https://doi.org/10.1016/j.geomorph.2012.05.007>
- Chen FX, Fang NF, Wang YX et al (2017) Biomarkers in sedimentary sequences: indicators to track sediment sources over decadal timescales. *Geomorphology* 278:1–11. <https://doi.org/10.1016/j.geomorph.2016.10.027>
- Chen Y, Wang Y, Yu K et al (2022) Occurrence characteristics and source appointment of polycyclic aromatic hydrocarbons and n-alkanes over the past 100 years in southwest China. *Sci Total Environ* 808:151905. <https://doi.org/10.1016/j.scitotenv.2021.151905>

- Cooper RJ, Pedentchouk N, Hiscock KM et al (2015) Apportioning sources of organic matter in streambed sediments: an integrated molecular and compound-specific stable isotope approach. *Sci Total Environ* 520:187–197. <https://doi.org/10.1016/j.scitotenv.2015.03.058>
- Cranwell PA (1981) Diagenesis of free and bound lipids in terrestrial detritus deposited in a lacustrine sediment. *Org Geochem* 3:79–89. [https://doi.org/10.1016/0146-6380\(81\)90002-4](https://doi.org/10.1016/0146-6380(81)90002-4)
- De Mastro F, Cocozza C, Brunetti G, Traversa A (2020) Chemical and spectroscopic investigation of different soil fractions as affected by soil management. *Appl Sci* 10. <https://doi.org/10.3390/app10072571>
- Desmet PJJ, Govers G (1996) A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *J Soil Water Conserv* 51:427–433
- ESDAC (2014) Soil Erodibility (K-Factor) High Resolution dataset for Europe. In: *Jt Res Cent Eur Comm*. <https://esdac.jrc.ec.europa.eu/content/soil-erodibility-k-factor-high-resolution-dataset-europe>. Accessed 7 Jun 2019
- ESDAC (2015a) Estimating the soil erosion cover-management factor at European scale. In: *Jt Res Cent Eur Comm*. <https://esdac.jrc.ec.europa.eu/content/cover-management-factor-c-factor-eu>. Accessed 19 Mar 2012
- ESDAC (2015b) Rainfall Erosivity in the EU and Switzerland (R-factor). In: *Eur Comm Jt Res Cent*. <https://esdac.jrc.ec.europa.eu/content/rainfall-erosivity-european-union-and-switzerland>. Accessed 7 Jun 2019
- ESRI (2017) ArcGIS Desktop
- Fang J, Wu F, Xiong Y et al (2014) Source characterization of sedimentary organic matter using molecular and stable carbon isotopic composition of n-alkanes and fatty acids in sediment core from Lake Dianchi, China. *Sci Total Environ* 473–474:410–421. <https://doi.org/10.1016/j.scitotenv.2013.10.066>
- Felde VJMNL, Schweizer SA, Biesgen D et al (2021) Wet sieving versus dry crushing: Soil microaggregates reveal different physical structure, bacterial diversity and organic matter composition in a clay gradient. *Eur J Soil Sci* 72:810–828. <https://doi.org/10.1111/ejss.13014>
- Ferraccioli F, Gerard F, Robinson C et al (2014) LiDAR based Digital Terrain Model (DTM) data for South West England.
- Ferrari AE, Ravnkov S, Larsen J et al (2015) Crop rotation and seasonal effects on fatty acid profiles of neutral and phospholipids extracted from no-till agricultural soils. *Soil Use Manag* 31:165–175. <https://doi.org/10.1111/sum.12165>
- Ficken KJ, Li B, Swain DL, Eglinton G (2000) An n-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. *Org Geochem* 31:745–749. [https://doi.org/10.1016/S0146-6380\(00\)00081-4](https://doi.org/10.1016/S0146-6380(00)00081-4)
- Gadiga BL, Martins AK (2015) The use of Revised Universal Soil Loss Equation (RUSLE) as a potential technique in mapping areas vulnerable to soil erosion in the Upper Yedzaram catchment of Mubi. *Int J Eng Res* V4:432–437. <https://doi.org/10.17577/ijertv4is030329>
- Geng J, Cheng S, Fang H et al (2019) Different molecular characterization of soil particulate fractions under N deposition in a subtropical forest. *Forests* 10:1–18. <https://doi.org/10.3390/f10100914>
- Glaser B, Zech W (2005) Reconstruction of climate and landscape changes in a high mountain lake catchment in the Gorkha Himal, Nepal during the Late Glacial and Holocene as deduced from radiocarbon and compound-specific stable isotope analysis of terrestrial, aquatic and microbi. *Org Geochem* 36:1086–1098. <https://doi.org/10.1016/j.orggeochem.2005.01.015>
- Glendell M, Jones R, Dungait JAJ et al (2018) Tracing of particulate organic C sources across the terrestrial-aquatic continuum, a case study at the catchment scale (Carminowe Creek, southwest England). *Sci Total Environ* 616:1077–1088. <https://doi.org/10.1016/j.scitotenv.2017.10.211>
- Gobin A, Jones R, Kirkby M et al (2004) Indicators for pan-European assessment and monitoring of soil erosion by water. *Environ Sci Policy* 7:25–38. <https://doi.org/10.1016/j.envsci.2003.09.004>
- Grabowski RC, Gurnell AM (2016) Diagnosing problems of fine sediment delivery and transfer in a lowland catchment. *Aquat Sci* 78:95–106. <https://doi.org/10.1007/s00027-015-0426-3>
- Guo J, Glendell M, Meersmans J et al (2020) Assessing branched tetraether lipids as tracers of soil organic carbon transport through the Carminowe Creek catchment (southwest England). *Biogeosciences* 17:3183–3201. <https://doi.org/10.5194/bg-17-3183-2020>
- Hijmans RJ (2020) Raster: Geographic Data Analysis and Modeling. R package
- Hirave P, Glendell M, Birkholz A, Alewell C (2020a) Compound-specific isotope analysis with nested sampling approach detects spatial and temporal variability in the sources of suspended sediments in a Scottish mesoscale catchment. *Sci Total Environ* 142916. <https://doi.org/10.1016/j.scitotenv.2020.142916>
- Hirave P, Wiesenberg GLB, Birkholz A, Alewell C (2020b) Understanding the effects of early degradation on isotopic tracers: implications for sediment source attribution using compound-specific isotope analysis (CSIA). *Biogeosci Dis* 1–18. <https://doi.org/10.5194/bg-2019-205>
- Krasa J, Dostal T, Jachymova B et al (2019) Soil erosion as a source of sediment and phosphorus in rivers and reservoirs – watershed analyses using WaTEM/SEDEM. *Environ Res* 171:470–483. <https://doi.org/10.1016/j.envres.2019.01.044>
- Lacey JP, Evrard O, Smith HG et al (2017) The challenges and opportunities of addressing particle size effects in sediment source fingerprinting: a review. *Earth-Science Rev* 169:85–103. <https://doi.org/10.1016/j.earscirev.2017.04.009>
- Lewis GP, Weigel AM, Duskin KM, Haney DC (2021) Wood abundance in urban and rural streams in northwestern South Carolina. *Hydrobiologia* 2. <https://doi.org/10.1007/s10750-021-04638-2>
- Liu X, Zou X, Cao M, Luo T (2020) Organic carbon storage and ¹⁴C apparent age of upland and riparian soils in a montane subtropical moist forest of southwestern China. *Forests* 11:1–19. <https://doi.org/10.3390/f11060645>
- Luo Y, Wang H, Meersmans J et al (2021) Modeling soil erosion between 1985 and 2014 in three watersheds on the carbonate-rock dominated Guizhou Plateau, SW China, using WaTEM/SEDEM. *Prog Phys Geogr* 45:53–81. <https://doi.org/10.1177/0309133320961274>
- Mayer S, Kühnel A, Burmeister J et al (2019) Controlling factors of organic carbon stocks in agricultural topsoils and subsoils of Bavaria. *Soil Tillage Res* 192:22–32. <https://doi.org/10.1016/j.still.2019.04.021>
- Meersmans J, Martin MP, De Ridder F et al (2012) A novel soil organic C model using climate, soil type and management data at the national scale in France. *Agron Sustain Dev* 32:873–888. <https://doi.org/10.1007/s13593-012-0085-x>
- Met Office (2021) UK climate averages. <https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-climate-averages/gbukbmeyr>. Accessed 14 Jan 2021
- Meyers PA (2003) Application of organic geochemistry to paleolimnological reconstruction: a summary of examples from the Laurentian Great Lakes. *Org Geochem* 34:261–289. [https://doi.org/10.1016/S0146-6380\(02\)00168-7](https://doi.org/10.1016/S0146-6380(02)00168-7)
- Nitzsche KN, Kleeberg A, Hoffmann C et al (2022) Kettle holes reflect the biogeochemical characteristics of their catchment area and the intensity of the element-specific input. *J Soils Sediments* 22:994–1009. <https://doi.org/10.1007/s11368-022-03145-8>
- Oliveira PTS, Nearing MA, Wendland E (2015) Orders of magnitude increase in soil erosion associated with land use change from native to cultivated vegetation in a Brazilian savannah environment. *Earth Surf Process Landforms* 40:1524–1532. <https://doi.org/10.1002/esp.3738>

- Owens PN (2020) Soil erosion and sediment dynamics in the Anthropocene: a review of human impacts during a period of rapid global environmental change. *J Soils Sediments* 20:4115–4143. <https://doi.org/10.1007/s11368-020-02815-9>
- Owens PN, Blake WH, Gaspar L et al (2016) Fingerprinting and tracing the sources of soils and sediments: Earth and ocean science, geoarchaeological, forensic, and human health applications. *Earth-Science Rev* 162:1–23. <https://doi.org/10.1016/j.earscirev.2016.08.012>
- Panagos P, Ballabio C, Borrelli P et al (2015) Rainfall erosivity in Europe. *Sci Total Environ* 511:801–814. <https://doi.org/10.1016/j.scitotenv.2015.01.008>
- Panagos P, Meusburger K, Ballabio C et al (2014) Soil erodibility in Europe: a high-resolution dataset based on LUCAS. *Sci Total Environ* 479–480:189–200. <https://doi.org/10.1016/j.scitotenv.2014.02.010>
- Pebesma EJ (2004) Multivariable geostatistics in S: the gstat package. *Comput Geosci* 30:683–691. <https://doi.org/10.1016/j.cageo.2004.03.012>
- Pebesma EJ, Bivand RS (2005) Classes and methods for spatial data in R: The sp package. *R News* 5 (2)
- Poesen J (2018) Soil erosion in the Anthropocene: research needs. *Earth Surf Process Landforms* 43:64–84. <https://doi.org/10.1002/esp.4250>
- Puttock A, Dungait JAJ, Macleod CJA et al (2014) Organic carbon from dryland soils. *J Geophys Res Biogeosciences* 119:2345–2357. <https://doi.org/10.1002/2014JG002635>
- Quenea K, Derenne S, Largeau C et al (2004) Variation in lipid relative abundance and composition among different particle size fractions of a forest soil. *Org Geochem* 35:1355–1370. <https://doi.org/10.1016/j.orggeochem.2004.03.010>
- Quénéá K, Largeau C, Derenne S et al (2006) Molecular and isotopic study of lipids in particle size fractions of a sandy cultivated soil (Cestas cultivation sequence, southwest France): Sources, degradation, and comparison with Cestas forest soil. *Org Geochem* 37:20–44. <https://doi.org/10.1016/j.orggeochem.2005.08.021>
- R Core Team (2020) R: A language and environment for statistical computing
- Ranjan RK, Routh J, Val Klump J, Ramanathan AL (2015) Sediment biomarker profiles trace organic matter input in the Pichavaram mangrove complex, southeastern India. *Mar Chem* 171:44–57. <https://doi.org/10.1016/j.marchem.2015.02.001>
- Renard KG, Foster GR, Weesies GA et al (1997) Predicting soil erosion by water: a guide to conservation planning with the revised universal soil loss equation (RUSLE). U.S. Dept. of Agriculture, Agricultural Research Service, Washington, D.C
- Rickson RJ (2014) Can control of soil erosion mitigate water pollution by sediments? *Sci Total Environ* 468–469:1187–1197. <https://doi.org/10.1016/j.scitotenv.2013.05.057>
- RStudio Team (2018) RStudio: Integrated Development for R
- Singh SN, Kumari B, Mishra S (2012) Microbial degradation of alkanes. In: *Microbial Degradation of Xenobiotics*. pp 439–469
- Starr GC, Lal R, Malone R et al (2000) Modeling soil carbon transported by water erosion processes. *Land Degrad Dev* 11:83–91. [https://doi.org/10.1002/\(SICI\)1099-145X\(200001/02\)11:1%3c83::AID-LDR370%3e3.0.CO;2-W](https://doi.org/10.1002/(SICI)1099-145X(200001/02)11:1%3c83::AID-LDR370%3e3.0.CO;2-W)
- Stock BC, Jackson AL, Ward EJ et al (2018) MixSIAR model description. 2018
- Stock BC, Semmens BX (2016) MixSIAR GUI User Manual. Version 3:1
- Stout SA (2020) Leaf wax n-alkanes in leaves, litter, and surface soil in a low diversity, temperate deciduous angiosperm forest, Central Missouri, USA. *Chem Ecol* 810–826. <https://doi.org/10.1080/02757540.2020.1789118>
- Torres T, Ortiz JE, Martín-Sánchez D et al (2014) The long Pleistocene record from the pego-oliva marshland (Alicante-Valencia, Spain). *Geol Soc Spec Publ* 388:429–452. <https://doi.org/10.1144/SP388.2>
- Trimble SW (1983) A sediment budget for Coon Creek basin in the Driftless Area, Wisconsin, 1853–1977. *Am J Sci* 283:454–474
- Valera CA, Pissarra TCT, Filho MVM et al (2019) The buffer capacity of riparian vegetation to control water quality in anthropogenic catchments from a legally protected area: a critical view over the Brazilian new forest code. *Water (Switzerland)* 11. <https://doi.org/10.3390/w11030549>
- Van Oost K, Govers G, Desmet P (2000) Evaluating the effects of changes in landscape structure on soil erosion by water and tillage. *Landsc Ecol* 15:577–589. <https://doi.org/10.1023/A:1008198215674>
- Van Oost K, Van Rompaey AJJ, Notebaert B et al (2006) WaTEM/SEDEM version 2006 Manual
- Van Rompaey AJJ, Govers G (2002) Data quality and model complexity for regional scale soil erosion prediction. *Int J Geogr Inf Sci* 16:663–680. <https://doi.org/10.1080/13658810210148561>
- Van Rompaey AJJ, Verstraeten G, van Oost K et al (2001) Modelling mean annual sediment yield using a distributed approach. *Earth Surf Process Landforms* 1236:1221–1236. <https://doi.org/10.1002/esp.275>
- Verstraeten G, Van Oost K, Van Rompaey A et al (2002) Evaluating an integrated approach to catchment management to reduce soil loss and sediment pollution through modelling. *Soil Use Manag* 18:386–394. <https://doi.org/10.1079/SUM2002150>
- Vogel HJ, Bartke S, Daedlow K et al (2018) A systemic approach for modeling soil functions. *Soil* 4:83–92. <https://doi.org/10.5194/soil-4-83-2018>
- Wang Y, Yang H, Zhang J et al (2015) Biomarker and stable carbon isotopic signatures for 100–200 year sediment record in the Chaihe catchment in southwest China. *Sci Total Environ* 502:266–275. <https://doi.org/10.1016/j.scitotenv.2014.09.017>
- Wiesmeier M, Urbanski L, Hobbey E et al (2019) Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales. *Geoderma* 333:149–162. <https://doi.org/10.1016/j.geoderma.2018.07.026>
- Wischmeier W, Smith D (1978) Predicting rainfall erosion losses: a guide to conservation planning. *Agricultural Handbook No. 537*
- Wohl E, Hall RO, Lininger KB et al (2017) Carbon dynamics of river corridors and the effects of human alterations. *Ecol Monogr* 87:379–409. <https://doi.org/10.1002/ecm.1261>
- Wu CL, Herrington SJ, Charry B et al (2021) Assessing the potential of riparian reforestation to facilitate watershed climate adaptation. *J Environ Manage* 277:111431. <https://doi.org/10.1016/j.jenvman.2020.111431>
- Yu M, Eglinton TI, Haghypour N et al (2019) Molecular isotopic insights into hydrodynamic controls on fluvial suspended particulate organic matter transport. *Geochim Cosmochim Acta* 262:78–91. <https://doi.org/10.1016/j.gca.2019.07.040>
- Zech M, Buggle B, Leiber K et al (2009) Reconstructing Quaternary vegetation history in the Carpathian Basin, SE-Europe, using n-alkane biomarkers as molecular fossils: Problems and possible solutions, potential and limitations. *Quat Sci J* 58:148–155. <https://doi.org/10.3285/eg.58.2.03>
- Zech M, Krause T, Meszner S, Faust D (2013) Incorrect when uncorrected: reconstructing vegetation history using n-alkane biomarkers in loess-paleosol sequences - A case study from the Saxonian loess region, Germany. *Quat Int* 296:108–116. <https://doi.org/10.1016/j.quaint.2012.01.023>
- Zhang X, Liu X, Zhang M, Dahlgren RA (2010) A review of vegetated buffers and a meta-analysis of their mitigation efficacy in reducing nonpoint source pollution. *J Environ Qual* 39:76–84. <https://doi.org/10.2134/jeq2008.0496>
- Zhang Y, Collins AL, McMillan S et al (2017) Fingerprinting source contributions to bed sediment-associated organic matter in the

headwater subcatchments of the River Itchen SAC, Hampshire, UK. *River Res Appl* 33:1515–1526. <https://doi.org/10.1002/rra.3172>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

C. Wiltshire¹  · M. Glendell² · T. W. Waine¹ · R. C. Grabowski¹ · J. Meersmans^{1,3}

¹ School of Water, Energy and Environment, Cranfield University, Bedford MK430AL, United Kingdom

² The James Hutton Institute, Craigiebuckler, AB158QH, Aberdeen, Scotland, United Kingdom

³ TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium