Human impact on long-term organic carbon export to rivers

Valentina Noacco1, Thorsten Wagener1,2, Fred Worrall3, Tim P. Burt4 and Nicholas J. K. Howden1,2

1Department of Civil Engineering, University of Bristol, Bristol, UK, 2Cabot Institute, University of Bristol, Bristol, UK, 3Department of Earth Sciences, Durham University, Durham, UK, 4Department of Geography, Durham University, Durham, UK

Abstract Anthropogenic landscape alterations have increased global carbon transported by rivers to oceans since preindustrial times. Few suitable observational data sets exist to distinguish different drivers of carbon increase, given that alterations only reveal their impact on fluvial dissolved organic carbon (DOC) over long time periods. We use the world’s longest record of DOC concentrations (130 years) to identify key drivers of DOC change in the Thames basin (UK). We show that 90% of the long-term rise in fluvial DOC is explained by increased urbanization, which released to the river 671 kt C over the entire period. This source of carbon is linked to rising population, due to increased sewage effluent. Soil disturbance from land use change explained shorter-term fluvial responses. The largest land use disturbance was during the Second World War, when almost half the grassland area in the catchment was converted into arable land, which released 45 kt C from soils to the river. Carbon that had built up in soils over decades was released to the river in only a few years. Our work suggests that widespread population growth may have a greater influence on fluvial DOC trends than previously thought.

1. Introduction

Global rivers transport, turn over, or store around 2.8 Pg of terrestrial organic carbon per year, a large fraction of the net terrestrial production [Cole et al., 2007; Battin et al., 2009a; Regnier et al., 2013]. Fluxes of carbon to world rivers have increased by 1.1 Pg C yr\(^{-1}\) since the preindustrial era (with a similar amount released from them), and previous studies suggest soils as the main source [Regnier et al., 2013; Leach et al., 2016]. These fluxes are a net C transfer between two of the world’s major carbon stores, the terrestrial biosphere and the ocean, so it is important to elucidate the physical mechanisms that underlie this significant change [Hope et al., 1994; Cole et al., 2007; Jansen et al., 2014] for informed decision making, policy development, and insight into future trends. However, there are currently no long-term observations of key drivers and responses, so long-term impacts of land use change on fluvial organic carbon concentrations and fluxes remain poorly understood [Wilson and Xenopoulos, 2009; Hering et al., 2015; Putro et al., 2016]. Research studies have typically used relatively short time series (usually post-1960) [Worrall et al., 2003; Monteith et al., 2007] to examine dissolved organic carbon (DOC) in basins in boreal climates [Larsen et al., 2011; Laidon et al., 2011, Couture et al., 2012; Ledesma et al., 2012; Tiwari et al., 2014] or those with organic-rich (e.g., peat) soils [Worrall et al., 2004; Worrall and Burt, 2005; Evans et al., 2006]. The focus of such work has been largely rural, often upland basins which are sparsely populated. In these cases, the primary interest has been aesthetic water quality [Mitchell, 1990], the formation of carcinogenic by-products during disinfection with chlorine [Hsu et al., 2001], or loss of organic C stocks from the northern peatlands [Evans et al., 2006; Laidon et al., 2011]. Several hypotheses have been proposed to explain these observed fluvial DOC increases as driven by changes in temperature [Freeman et al., 2001; Worrall et al., 2004; Tian et al., 2013], atmospheric CO₂ [Freeman et al., 2004], atmospheric sulphur deposition [Clark et al., 2005; Monteith et al., 2007], nitrogen deposition [Findlay, 2005], hydrology and precipitation [McDowell and Likens, 1988; Hongve et al., 2004; Erlindsson et al., 2008], land use [Garnett et al., 2000; Anderson et al., 2013], or land management [Yallop and Clutterbuck, 2009]. In contrast, there has been little focus to quantify fluvial organic C transfers from terrestrial ecosystems to the oceans in lowland basins that host both intensive agricultural activity, and large and growing human (predominantly urban) populations, mainly due to the absence of long-term data to compare the relative influence of different drivers (past and present) of natural and anthropogenic activity in such areas.
Here we investigate the main drivers for the increase in fluvial DOC (Figure 1) in the temperate, lowland, mineral soil-dominated Thames basin, which has a significant urban population, and estimate their relative contributions. We develop a model linking basin properties and DOC release to the river by adapting a UK nation-scale DOC export model [Worrall et al., 2012] and combine it with a model of the effects of land use and land use change (LU and LUC) on soil organic carbon (SOC) release [Bell et al., 2011] to estimate the relative influence of drivers of land use, climate, and population on DOC dynamics for the Thames basin since 1884.

2. Study Site

The River Thames basin (Figure 2) upstream of London, England, drains a 9948 km$^2$ temperate, lowland, mineral soil-dominated catchment. It is the second longest river in the UK, with a total length of 354 km [Marsh and Hannaford, 2008]. It has a large urban population, which has grown fourfold since the 1880s. The gauged Thames basin contains many major urban centers, including Swindon, Oxford, Slough, Reading, and Maidenhead (Figure 2b). It provides two thirds of London’s drinking water [Environment Agency, 2009]. Moreover, it has many sewage treatment works (STWs), related to its high human population density (~960 people km$^{-2}$) [Merrett, 2007]. Tertiary wastewater treatment has been installed at the 36 largest STWs (serving approximately 2.7 million people) upstream of the tidal limit since 2003 [Kinniburgh and Barnett, 2009]. The Thames basin population is expected to increase still further in the near future [Environment Agency, 2009]. The continuing rise in the population of the basin will increase pressure on drinking water supplies, wastewater treatment, water quality, and river ecology even further [Evans et al., 2003; Neal and Jarvie, 2005]. Moreover, this pressure will probably worsen with projected climate change scenarios, which predict decreasing river flows and increasing water temperatures [Johnson et al., 2009]. In spite of the high population density, the Thames basin upstream of London is predominantly rural (87.7% land area in 2007 [Centre for Ecology and Hydrology (CEH), 2011]). The Thames basin is mostly underlain by Cretaceous Chalk, with areas of limestone, mudstones, sandstones, and Oxford clay [Howden et al., 2011].

3. Model Setup

We predicted the riverine DOC concentration trend in the Thames River between 1884 and 2005 by estimating the contribution of natural and anthropogenic drivers. We achieved this by adapting a steady state model that estimated DOC export from basin characteristics, land use, sewage effluents, and net losses at the UK
scale [Worrall et al., 2012], and combining it with a model that captured the effect of land use transitions on the release or sink of SOC [Bell et al., 2011].

3.1. SOC Stock Model

We adapted the parsimonious UK SOC model [Bell et al., 2011] to estimate the SOC stock for the Thames basin. We used the Thames basin annual land use history from 1884 to 2005 [Howden et al., 2011, 2013], together with typical basin-specific land use SOC concentrations, and first-order rate constants, which take into account the change in SOC concentration following land use change. The model accounts for three contributions: the SOC stock in areas in the basin that have not undergone land use changes since the beginning of the record (equation (1)), the change in SOC stock due to land use transitions (loss or gain) (equation (2)), and the loss of SOC stock due to increase in temperature (equation (3)).

The carbon stored in areas that have not changed their land use since the beginning of the record is

$$\text{SOC}_{\text{stock}} = K \sum_i A_i d(C_{o_i} \rho_i)$$

where $K$ is a factor to equalize units, $A_i$ is the area under land use $i$ (km$^2$), $d$ is the depth of the soil layer considered (cm), $C_{o_i}$ is the organic C content of the soil for a given land use ($\%$), and $\rho_i$ is the soil bulk density of a given land use (g cm$^{-3}$).

Following Bell et al. [2011], we assumed that SOC transitions after land use change follow first-order rate kinetics:

$$\Delta \text{SOC}_{\text{stock}} = K \sum_i \sum_j A_{ij} d(C_{o_i} \rho_i - C_{o_j} \rho_j) e^{-\lambda(t-t_i)}$$

where $\Delta \text{SOC}_{\text{stock}}$ is C stock variation after land use change in the basin soils in year $t > t_i$ (where $t_i$ is the
year when the land use change occurred) (tons C), $A_{xy}$ is the area changed from land use $j$ to $i$ (km$^2$) (which is done for each combination of land use transition), and $\lambda$ is the time constant which accounts for the time it takes SOC concentration to adjust to a new concentration following land use change (yr$^{-1}$). Equation (2) is the sum of SOC stock each year in the soils of the Thames basin at depth $d$ after all land use transitions have been taken into account. Equation (2) accounts for both increases and decreases in SOC stock (e.g., a conversion from temporary grassland to woodland would cause an increase in SOC stock over time, while a conversion of permanent grassland to arable land would cause a decrease of SOC stock over time).

We modified the model to account for SOC changes due to temperature increase by using the study of Smith et al. [2007]. We estimated the increase in soil respiration rate per 10°K temperature increase, $Q_{10}$ [Kutsch et al., 2007], from

$$R_f = R_{t0} Q_{10}^{\frac{T_{t0} - T_f}{10}}$$

(3)

where $R_f$ is the soil respiration rate at Kelvin temperature $T_f$. Soil respiration rates were considered in equilibrium with the mean annual temperature. Smith et al. [2007] suggest that the soil respiration rate doubles with a temperature increase of 10°K (i.e., $Q_{10} = 2$), which gives a UK total soil respiration rate between 1.3 and 2.6 Tg C yr$^{-1}$, equivalent to between 0.053 and 0.106 Tg C yr$^{-1}$ scaled to the Thames basin area. To account for the contribution of temperature increase, we assumed that the SOC values were in equilibrium with the respiration rates of Smith et al. [2007] in the year before the start of the study ($t_0$). This scenario used the range of respiration rate proposed above as $R_{t0}$, and estimated the range of $R_f$ for every year with equation (3).

We then subtracted the simulated respiration rates from the SOC stock to give a joint estimate of the SOC stock after land use change and the effect of increasing temperature (i.e., loss of SOC due to increasing temperature over the period due to increased soil respiration and therefore loss of C to the atmosphere) (equation (4)).

$$\text{Total SOC}_{stock} = \text{SOC}_{stock} + \Delta\text{SOC}_{stock} = R_f$$

(4)

We used the model in an uncertainty framework by randomly resampling 100,000 values from the observed SOC concentrations from the range below their 75th percentile. Given that the SOC concentration distributions are right-skewed for the four land uses reported in the databases available (arable, temporary grassland, permanent grassland, and woodland) (Table S1 in the supporting information and Figure 3), our procedure has two advantages. First, no predefined distribution is superimposed to the data, as the data are randomly sampled from the measured SOC concentrations reported in the databases, thus preserving the original distributions. Second, by sampling below, the 75th percentile outliers are ignored and the mass of the distributions is well represented. This procedure has the disadvantage that since we are sampling only from observed values, SOC concentrations that have not been measured, but which are equally likely, are not represented. However, given the extensive number of observed values, this should provide a good coverage of the most likely SOC concentrations. Since we could not access the sampled measurement values for SOC concentrations in urban areas, we used the SOC concentration in the urban area of Coventry from the study of Rawlins et al. [2008] (Table S1). The data have similar mean and median, which suggests that the distribution is symmetric; therefore, we approximated the distribution with a normal, and we sampled from it. For land use transition constants and soil bulk densities, their median values were chosen to avoid unrealistic SOC content due to stochastic selection of both SOC concentration and bulk densities from the top/bottom of their ranges.

Annual SOC fluxes were calculated as the interannual change in the SOC stock:

$$\text{SOC}_{flux} = \text{Total SOC}_{stock} - \text{Total SOC}_{stock}$$

(5)

where SOC$$_{flux}$$ is the flux of C from basin soils in year $t$ (tons C); thus, SOC$$_{flux}$$ < 0 represents C sequestration and SOC$$_{flux}$$ > 0 represents release of SOC to the atmosphere and water systems [Bell et al., 2011].

The median values (with the upper and lower quartiles) of the 100,000 simulations for SOC stock and flux are plotted in Figure 4. The overall contribution from soils is roughly in balance over the period, even though we acknowledge that due to the uncertainty of the data, it could be either a net sink or source. A UK nation-scale study has shown that UK soils acted as an overall C sink over the period of 1925–2007, mainly due to the creation of woodland and the conversion of arable land into permanent grassland [Bell et al., 2011]. Future efforts
Figure 3. Empirical density and cumulative distributions of soil organic carbon (SOC) content and soil bulk density (SBD) for the four land use types in the SOC stock model. Available samples for arable (SOC: $n = 592$; SBD: $n = 242$), temporary grassland (SOC: $n = 205$; SBD: $n = 65$), permanent grassland (SOC: $n = 367$; SBD: $n = 111$), and woodland (SOC: $n = 82$; SBD: $n = 65$).
will need to be directed toward further sampling of SOC concentrations of UK soils and on long-term field studies to measure SOC change in order to reduce this uncertainty, given its potential implications.

### 3.2. Dissolved Organic Carbon Export Model

The DOC export model developed by Worrall et al. [2012] explored what are the physically meaningful catchment characteristics controlling DOC fluxes from various UK catchments. An informed regression based on 169 UK catchments was obtained by selecting the variables which were both statistically significant ($p < 0.05$) and physically meaningful. Moreover, only variables that could be mapped or extrapolated across Great Britain were selected. The variables tested, which were not found to be statistically significant, are reported in Table S2. The model obtained was

$$\text{DOC}_{\text{flux}} = 2.6 \text{Mineral} + 3.4 \text{OrgMin} + 9.2 \text{Organic} + 6.7 \text{Urban} + 2.4 \text{Grass} - 2.7 \text{Area}$$

where $\text{DOC}_{\text{flux}}$ is the average DOC flux ($t \text{ C yr}^{-1}$) and Mineral, OrgMin, Organic, Urban, and Grass are the area of mineral, organo-mineral and organic soils, and area under urban development and grassland in the Thames Basin.

![Figure 4](https://example.com/fig4.png)

**Figure 4.** Modeled SOC flux in the Thames Basin. (a) Modeled soil organic carbon fluxes resulting from land use change and temperature increase between 1884 and 2005 estimated as the interannual variation of soil organic carbon stock. A positive flux represents a loss of carbon from the soil to the atmosphere and water system; a negative flux represents carbon sequestration in the soil. (b) Cumulative soil organic carbon flux between 1884 and 2005. The solid lines represent the median values of the ensemble estimates, and the dashed lines represent the interquartile range.
organic, organo-mineral and mineral soils, point-source discharges from wastewater treatment, and occasional net releases following land use change (e.g., a conversion from permanent grassland to arable land leads to decreased C storage within that land proportion). The main carbon losses occur due to in-stream processing, for which we use catchment area as a proxy; mineralization from soils, for which we accounted using temperature change in the land use change model; and occasional sinks/storages that occur following certain land use changes (e.g., a conversion from temporary grassland to woodland leads to increased C storage within that land proportion). Our final model was

\[ \text{DOC}_{\text{flux}} = 2.6 \text{Mineral} + 3.4 \text{OrgMin} + 9.2 \text{Organic} + 6.7 \text{Urban} + \alpha \text{SOC flux} - 2.7 \text{Area} \]  

where the relative presence of mineral and organo-mineral and organic soil types (86%, 12%, and 2%, respectively) in the Thames basin is a proxy for the natural carbon baseline in the river coming from soils. The \( \alpha \) is the export coefficient of carbon from soils to surface runoff. The urban land use type is used for two purposes. First, the transient term Urban in equation (7) allows for the increase in urbanization in the Thames catchment throughout the period; this is a proxy for the increase in population, which brings an increase in sewage effluent discharge, combined sewer overflows (CSOs), impervious surface runoff, industrial wastewater, etc. [Tian et al., 2012]. Second, urban area is included in the synthesis of all land use changes in the term SOC flux in equation (7), which is the annual SOC flux from the basin (from equation (5)). In this context we consider the amount of C that can be stored in urban soils (e.g., the conversion of permanent grassland into urban area is a loss of C from soils, and so reduced C storage in the catchment soils can lead to increased DOC export to the river). Grassland was originally used as a proxy estimate of all land use states in the catchments over the period of record considered [Worrall et al., 2012]. We substituted it with SOC flux, because it provides a representation of land use dynamics and transitions, as this is more relevant to a 120 year study. This substitution could be done because the grassland predictor was not correlated to the other predictors in the national multiple linear regression, and therefore, it was independent. The \( \alpha = 0.0165 \) was derived from the export coefficient from soils to surface and groundwater pathways (3%), and the surface runoff partition for the Thames basin (55%) [Howden et al., 2011]. Therefore, 1.65% of the C lost from soils leaches to the river system. A summary of the parameters used in the model is provided in Table 1. The parameters most influential are those for the Mineral and the Area terms, as these are the terms that contribute the most to the output. The SOC flux term can also be negative, which means a decline in the baseline C export from the catchment. For example, a change in land use from arable to grassland leads to negative SOC flux, i.e., increased C sequestration in soil, and so less C export from soils compared to the baseline conditions when that area was under arable land use. The Area term in equation (7) is used as a proxy for the in-stream removal of DOC along the catchment, given that the area of the basin is used as a proxy for the length of the river, and therefore for the in-stream residence time of DOC, and so a proxy for organic matter decomposition, photodecomposition, flocculation, sequestration in sediments, etc.

DOC flux is estimated as the sum of the ensembles of 100,000 Monte Carlo estimates of each model term (i.e., input from soils, in-stream removal, and urbanization) and the ensemble of the 100,000 estimates of SOC flux (see the preceding section). The estimated DOC concentration was then obtained by dividing the estimated DOC flux by the average annual flow over the period. This was assumed acceptable given that annual mean streamflow, even though is predicted to decrease in the future [Johnson et al., 2009; Jin et al., 2012], did not change significantly over the period of 1884 to 2013 (p > 0.5) (Figure 5c).

| Table 1. Key Parameters Used for DOC Export Model | Export Coefficient | Units |
| Urban | 6.7 | t C km\(^{-2}\) yr\(^{-1}\) |
| SOC flux\(^a\) | 0.0165 | – |
| Mineral soil\(^a\) | 2.6 | t C km\(^{-2}\) yr\(^{-1}\) |
| Organo-mineral soil\(^a\) | 3.4 | t C km\(^{-2}\) yr\(^{-1}\) |
| Organic soil\(^a\) | 9.2 | t C km\(^{-2}\) yr\(^{-1}\) |
| Area\(^a\) | –2.7 | t C km\(^{-2}\) yr\(^{-1}\) |

\(^a\)From Worrall et al. [2012], \(^b\)From Howden et al. [2011].
4. Data Description

4.1. DOC Concentration Data

Water color data and DOC concentration measurements (>20,000 samples) were obtained by statutory authorities between 1883 and 2014 (Figure 1 and Table 2) for the Thames at Hampton (51.42°N, 0.37°W) and at Teddington (51.43°N, 0.33°W), respectively. A review of the methods for measuring color (1883 to 1990) and DOC (1990 to 2014) and the calibration of DOC from color measurements is provided in the following section.

Water color data between 1883 and 1974 were obtained by examining water (clarified by filtration if necessary) by using a tube 0.6 m long and by comparing it to a solution of potassium dichromate and cobalt sulphate (the so-called Burgess method), which has a yellow-brown color, in a second tube, as long as the first [Burgess, 1902]. Enough of this solution is poured in the tube with distilled water to match the color of the examined water and then reported as ratio of brown to blue, which stands for millimeter of the Burgess solution poured to match the color of the examined water [Thresh et al., 1943]. Water color data between 1974 and 1990 were measured by the Environment Agency which reported water color for samples in Hazen units (equivalent to the platinum concentration in a standard solution of platinum/cobalt chloride salts of the same absorbance/color [Department of the Environment, 1972; Worrall and Burt, 2007]), measured with absorptiometer (601 filter or 430 μm) [Simpson, 1980]. By examining a large number of samples of water, Hazen units of color were found to correspond very closely to 2.2 Burgess units [Thresh et al., 1943]. Organic carbon
between 1883 and 1905 was obtained by measuring the amount of oxygen taken up by water from an acid solution of potassium permanganate, while the amount of carbon in the organic matter present in the water residue was determined with the eudiometric method devised by Frankland [Thresh, 1904]. DOC measurements between 1990 and 2014 were defined within the establishment of the Harmonised Monitoring Scheme (HMS) [Department of the Environment, 1972; Simpson, 1980].

The calibration between water color and DOC was performed for the water samples collected once per week for the water years between 1899 and 1905 (correlation = 0.81, \( p < 0.001 \)).

\[
\text{DOC} = 0.0955 \text{Color} + 0.2103 \quad R^2 = 0.66, n = 194
\]  

This calibration was then used to estimate the time series of DOC from water color data for the period between 1883 and 1990. It should be noted that this study has assumed that equation (8) was stationary with time. However, Worrall and Burt [2010] have demonstrated that DOC-color relationships have shifted across the UK over periods since 1974. However, in this study the estimates of equation (8) were compared with independent DOC data collected during the early period by the Grand Junction Water Company (1883 to 1905). We further collated these data with DOC concentration data from Thames water (1990 to 1998) and Environment Agency (EA) (1998 to 2014). DOC concentrations from these various sources were consistent (Figure 1). Additional analysis to explore the consequences of the use of different analytical techniques is presented in section 6.2.

### 4.2. Hydroclimate Data

Average monthly temperature and total monthly rainfall [Burt and Howden, 2011] data are available for the Radcliffe Observatory, Oxford, from 1815 (Figures 5a and 5b). Mean annual temperature and rainfall for the catchment at Oxford, which is centrally located within the basin, are 10.1°C and 652.7 mm (1883–2013, standard deviation of 0.7°C and 114.1 mm), respectively. Continuous river flow records were available for Teddington Weir from January 1883 (Figure 5c), with mean annual flow of 65.5 m\(^3\) s\(^{-1}\) (1883–2013, standard deviation of 26.7 m\(^3\) s\(^{-1}\)) (National River Flow Archive: http://www.ceh.ac.uk/data/nrfa/).

### 4.3. Data Used for SOC Stock Model

#### 4.3.1. Land Use Data

Land use data comprised five categories: arable, temporary grassland, permanent grassland, woodland, and urban. These categories are consistent with those used by other national-scale SOC stock-model studies [Bradley et al., 2005; Smith et al., 2010; Bell et al., 2011]. In our classification arable lands include crops and bare fallow; temporary grassland includes area under grassland arable rotation for less than 5 years; permanent grassland includes rough grazing; woodlands include broadleaf-deciduous woodland, coniferous woodland, and scrub/orchard; and urban includes urban and suburban. Woodland area was collated by linearly interpolating data from Crooks and Davies [2001] for the period between 1870 and 1990 and from the Centre for Ecology and Hydrology’s Land Cover Map 2007 (LCM 2007) [CEH, 2011]. Urban area was collated by linearly interpolating data from Crooks and Davies [2001] for the period between 1870 and 1938, and from LCM 2007. Urban area was then validated with the population data for the basin which confirmed the upward linear trend. Agricultural (i.e., arable, temporary, and permanent grasslands) data collated by Howden et al. [2011] were used, which originally come from parish records, and catchment Agricultural Census [Ministry of Agriculture, 1986]. The data for these three land use categories are assumed to be a representative sample of the relative importance of each of these land use types. Therefore, allowing for 5% of the area of the Thames catchment under a land use not included in our five categories (e.g., mountain, heath, bog, and

### Table 2. Data Provenance

<table>
<thead>
<tr>
<th>Period</th>
<th>Variable</th>
<th>Measuring Body</th>
<th>Sampling Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1883–1905</td>
<td>Color</td>
<td>Burgess Grand Junction Water Company</td>
<td>Daily</td>
</tr>
<tr>
<td>1883–1905</td>
<td>OC</td>
<td>Parts per 100,000 Grand Junction Water Company</td>
<td>Weekly</td>
</tr>
<tr>
<td>1974–1988</td>
<td>Color</td>
<td>Hazen Thames Water Authority</td>
<td>Monthly</td>
</tr>
<tr>
<td>1990–1998</td>
<td>DOC</td>
<td>mg L(^{-1}) Thames Water</td>
<td>Approximately monthly</td>
</tr>
<tr>
<td>1998–2014</td>
<td>DOC</td>
<td>mg L(^{-1}) Environment Agency</td>
<td>Approximately monthly</td>
</tr>
</tbody>
</table>
inland water), we used the proportion of each of these agricultural areas to partition the residual area (95% of the Thames basin area minus urban and woodland) between arable, temporary, and permanent grasslands.

### 4.3.2. SOC Concentration Data

SOC concentrations for soils under arable, temporary grassland, permanent grassland, and woodland were aggregated from seven databases to give a total of 1249 soil samples which considers an extensive range of circumstances over which SOC was sampled in the Thames basin [Bell et al., 2011]: National Soils Inventory (NSI) 1984 [Loveland, 1990], the National Soils Inventory (NSI) 2001 [Bellamy et al., 2005], the Countryside Survey (CSS) 1978 [Black et al., 2002], the Countryside Survey (CSS) 1998 [Haines-Young et al., 2000], the Representative Soil Sampling Scheme [Webb et al., 2001], and English Nature Woodland data (EW Woodland) 2001 (Figure 3 and Table S1). All SOC concentrations refer to the top 15 cm of soil, which is most affected by changes in land use and management [Kimble et al., 2000; Poeplau and Don, 2013]. Since none of the databases considered included measurements of SOC content in urban areas, and up to now there are no measurements relating to the Thames basin, it was considered the best approximation to utilize the estimates from Rawlins et al. [2008], as they relate to the topsoil (0–15 cm depth) for the urban area of Coventry (Table S1).

### 4.3.3. Soil Bulk Density Data

Soil bulk density was computed from the amalgamated databases for SOC concentrations (see above) that also reported bulk density (i.e., NSI 1984, NSI 2001, and EW Woodland 2001). Contrary to SOC, we used the median value given the low variability of bulk density (Table 3). As none of the databases used reported bulk density measurements for an urban area, the median value for arable land was used.

### 4.3.4. First-Order Rate Constants

First-order rate constants for each land use transition account for the time it takes for SOC concentration to adjust to a new concentration following land use change. Values used come from synthesis of the literature review performed by Bell et al. [2011] and are given in Table S3. When the half-lives were known, the first-order rate constants were calculated by using equation (9)

\[
\lambda = \ln(2) \frac{1}{t_{1/2}}
\]

where \( \lambda \) is the first-order rate constant (yr\(^{-1}\)), while \( t_{1/2} \) is the estimated half-life (yr).

Otherwise, half-lives were estimated by using information on SOC concentrations at different time intervals or other qualitative information on time needed to reach new equilibrium. Transitions to or from temporary and permanent grasslands were assumed to have the same first-order rate constants due to a lack of distinction between the two in the literature sources. Due to a lack of specific information on first-order rate constants for transition into urban, it was assumed that the transition would be instant, given the removal of soil during construction. No transition out of urban or woodland was reported in the Thames basin. The only transition into woodland reported was from arable; other transitions into woodland were assumed to be the same. As an example, the time it takes for soils to adjust to a new SOC concentration after conversion of permanent grassland to arable land is around 20 years, with a half-life of 2.5 years (\( \lambda = 0.28 \)). While for the inverse land use transition it takes around 120 years, with a half-life of 17.3 years (\( \lambda = 0.04 \)).

### 4.3.5. Land Use Transitions

Land use changes were attributed by the method recommended by Bell et al. [2011]: concurrent increases and decreases of similar magnitude in two land uses were paired (e.g., permanent grassland and arable land in the 1940s); otherwise, the land use transitions were ascribed following the most likely transitions as detailed by Adger and Subak [1996] and Adger et al. [1992] (Table S4). For each year the model considers all the transitions from one land use to another, so, for example, in 1940, the area under permanent grassland was converted into arable land, while temporary grassland was converted into arable, urban, and woodland (following the preferential directions presented in Table S4).
4.4. Data Used to Build the UK DOC Export Model

The DOC and water color (converted into DOC by using the calibration method developed by Worrrall and Burt [2007]) data used in the UK DOC export model come from the HMS, the EA, and the Scottish Environmental Protection Agency between 2001 and 2007. The catchments considered have an annual average flow at the tidal limit higher than 2 m$^3$ s$^{-1}$, an area bigger than 40 km$^2$ and at least one sample per month. The catchment properties investigated were soil type (classified into mineral, organo-mineral, and organic), catchment area (estimated up to the monitoring point from the CEH Wallingford digital terrain model), and land use (classified into arable, grassland, and urban based on the June Agricultural Census for 2004 [Department for Environment Food and Rural Affairs et al., 2013]).

5. Results

5.1. DOC Trends Over Time

The long-term observations suggest a positive trend in mean DOC concentrations, which doubled between 1884 (1.7 mg C L$^{-1}$) and 2014 (4.3 mg C L$^{-1}$) (Figure 1). Annual mean DOC concentration was stable prior to 1912, but increased monotonically until 1938, rising to a peak of 6.7 mg C L$^{-1}$ in 1941 and then declining to 2.8 mg C L$^{-1}$ in 1960. After two minor peaks in the 1960s and 1970s, DOC increased to its highest annual peak in 1995 (7.7 mg C L$^{-1}$), and then decreased during the last decade of the record.

5.2. DOC Load Contributions

Since 1881, urban area has increased by a factor of 2.5, and there were massive land conversions from permanent pasture into arable production during World War (WW) II [Howden et al., 2011, 2013] (Figure 2c). We estimate a net release to the river of 824.4 (275.8, 1374.5) kt C (interquartile range, hereafter) for Environment Food and Rural Affairs et al. (classifying up to the monitoring point from the CEH Wallingford digital terrain model), and land use (classified into arable, grassland, and urban based on the June Agricultural Census for 2004 [Department for Environment Food and Rural Affairs et al., 2013]).

5.3. Relative Importance of DOC Contributions

The baseline C release rate remains constant over time, but its relative importance decreases as other sources increase. LU and LUC are net sinks most years, of up to 30%, although they account for >50% of net annual releases during the 1940s (Figure 6a). Urban sources contributed 72% of total organic C inputs in 1884, rising to 87% in 2005; the only exception was during WWII when the impact of land use change reduced the proportional effect of urban inputs to as low as 34% in 1943. We estimate that these C-load components translate into a baseline DOC concentration of around 0.6 mg C L$^{-1}$, moderated by release/uptake of soil DOC by land use and land management of between −1 mg C L$^{-1}$ (2005) and...
4.7 mg C L\(^{-1}\) (1943). Transient DOC contributions from increasing urban area increased from 1.5 mg C L\(^{-1}\) (1884) to 4.1 mg C L\(^{-1}\) (2005), with 90% of the long-term trend in DOC concentrations being explained by the rise in C inputs from urban expansion (\(p < 0.001\)).

LU and LUC were a cumulative sink of 32.4 kt C (−0.6 kt C yr\(^{-1}\)) from 1884 to 1939. In 1940, large-scale conversion of permanent pasture into arable production resulted in a cumulative net source of 44.6 kt C by 1949, an average release of 4.5 kt yr\(^{-1}\). After 1950, LU and LUC returned to being a net sink (−0.3 kt C yr\(^{-1}\)), so by 2005 the net impact of LU and LUC over the whole period was roughly zero. Our model of LU and LUC also included the impact of changes in mean annual air temperature, since temperature controls the decomposition of organic matter in soils and thus DOC export [Kätterer et al., 1998]. Temperature increased at an average rate of 0.01°C (± 0.001) yr\(^{-1}\) (\(p < 0.001\)) between 1884 and 2005 (Figure 5a), and overall decreased SOC stocks by an average of 0.2%, an insignificant release of fluvial C compared to other drivers.

For the majority of the twentieth century, LU and LUC were small net C sinks, but they showed the potential for large C releases over short periods (a few years, i.e., <10 years) following large-scale land conversions, as occurred in the early 1940s. The land use change in the 1940s is evidence of how just one major land use conversion can release to rivers, and part of it in turn to the atmosphere, carbon accumulated in soils for hundreds of years [Butman et al., 2014].

Figure 6. Estimated DOC load and concentration (1884–2005 water years). (a) Relative importance of the contributions to DOC from the pre-1880 baseline, urbanization and land use change, and the total contribution. (b) Comparison between estimated and observed DOC concentrations and ranges. The blue dots show the observed annual average DOC concentration, the light blue shaded area shows the smoothed range of observed DOC concentration, the solid black line is the estimated annual DOC concentration, and the grey shaded area is the interquartile range of the estimated DOC concentration. (c) Prediction error between estimated and observed DOC.
6. Discussion

6.1. Capturing Long-Term DOC Drivers in a Highly Human-Impacted Basin

Human activities have greatly impacted the natural environment in the last century. These activities include the intensification of agricultural practices (e.g., extensive ploughing of grassland into arable land and the introduction of widespread mechanization and land drainage [Howden et al., 2010, 2013; Marsh and Harvey, 2012]) and the increase of urbanization and population for over a century. All these activities have repercussions on the carbon export to rivers [Regnier et al., 2013]. The Thames basin, whose long-term records date back to the late nineteenth century, allows us to capture the effect of these changes on natural fluvial DOC levels and to understand how humans have altered the pre-1880 baseline of DOC. The DOC concentration in the Thames catchment was around 1.8 mg C L\(^{-1}\) before 1900, which is the DOC concentration preceding the intensification of agriculture and major land use changes. During the last century, DOC concentration has increased to levels up to 6 times higher, with peaks reaching 12 mg C L\(^{-1}\) since the late 1980s. The expansion of urban area in the basin, which meant increased point sources of organic carbon from sewage effluent, released 671 kt C over the study period. The increase in agricultural productivity during WWI released 5 kt C into the river, but by far the biggest release due to land use change was during WWII, when the large-scale ploughing of grassland released to the river 45 kt of C, 5% of the total carbon released to the river over the entire 120 year period. In the early 1990s 5 kt of C were released due to the conversion of permanent grassland into temporary grassland.

The Thames basin has not experienced an increase in streamflow or precipitation over the study period, while mean annual temperature, which increased by 1.3°C (Figure 5), did not significantly increase DOC export to the river. This is in agreement with the study of Tian et al. [2013], who found that a variation in the annual mean temperature of at least 5°C was needed for a significant influence on DOC export.

Most studies concerned with the attribution of DOC increase to different drivers have been limited by the insufficient length of the available records. The exceptional length of our observational record enabled us to elucidate the long-term drivers for the increase in riverine DOC concentration in a basin that has undergone important historical changes. These changes, both gradual and abrupt, are not peculiar to the Thames basin: large-scale dramatic land use conversions, such as shifts in agriculture and urbanization, have occurred [Tian et al., 2015] and are still occurring [Davis et al., 2015] worldwide. The increase in urban area, which is a consequence of population increase, has been shown to be the main driver, contributing to more than 80% of the rise in DOC. This could be due to the increase in sewage effluent discharge, wastewater leaks, impervious surface runoff, or combined sewer overflows (CSOs) [Tian et al., 2012], which increased fluvial DOC levels, notwithstanding the upgrade of the sewerage infrastructure over the period and the improvement of the sewage treatment after the implementation of the Urban Waste Water Treatment Directive (UWWTD) in the 2000s [European Union, 1991]. UWWTD is expected to have reduced fluvial DOC, by imposing stricter levels of organic matter in sewage effluent discharges from STWs. But even with better treatment, DOC concentrations in effluent discharges are usually much higher than in receiving waters [Westerhoff and Anning, 2000; Sickman et al., 2007]. The implications of our findings are not restricted to the Thames basin, as half the global population is living in urban areas, and people moving from rural areas to cities is continuing [United Nations, 2008]. Our study highlights the important role urban areas will play in the carbon export to rivers and oceans in the future, given that 70% of the global population is expected to live in urban areas by 2050 [United Nations, 2008].

6.2. Statistical Analysis of the Impact of Change in Analytical Techniques

Different methods for measuring water color and DOC have been used over the period (1884–1974: DOC calibrated from Burgess units of color, 1975–1990: DOC calibrated from Hazen units of color, and 1991–2005: DOC concentration). We analyzed whether the differences in the annual DOC mean in the three periods were statistically significant by using a multiple linear regression model. First, we estimated a model where as predicted variable we used a factor with three levels, which indicates in which period we are. The effect of analytical technique on the mean of DOC was 1 mg C L\(^{-1}\) higher in the second period than in the first, while in the third it was 2.2 mg C L\(^{-1}\) higher than in the first (in both cases \(p < 0.005\)). We then estimated a model where as predicted variables we added also urban area and SOC flux (which, as we showed, explain the DOC trend). Here the effect of analytical technique on the mean of DOC in the second period was not significantly
In spite of the limitations presented above, the model, which was obtained without any optimization or different types of catchment, includes a loss term which is the average loss term across all sources (including urban) and across a range of different periods. Equation (6) is based upon the actual observations of over 160 catchments and their behavior, and it effect been accounted for by the manner in which DOC concentration prediction being affected by an enhanced lability of DOC from sewage sources has in 2000. Sewage effluents are also a source of more labile DOC [Painter and Viney, 1959; Tzortziou et al., 2015]). Therefore, we would expect a likely underestimation of the amount of observed DOC in the river pre-1990, due to the calibration period used. This would result in a greater increase in the observed fluvial DOC over the period when DOC is calibrated from color (1884–1990), as the increasing contribution of less colored carbon with increasing urbanization is not fully accounted for. This would partly reduce the observed steep increase in the observed DOC in the 1990s. Nonetheless, it is worth mentioning that such peaks have been observed before (DOC level during WWII is comparable to the peak in the 1990s) and there is consistency in DOC concentrations over the period when the observed measurements changed from color to DOC. Moreover, the DOC increase in the 1990s could be a result of the implementation of nitrate vulnerable zones (NVZs), which put limits to the application rates of manure and fertilizers, to their timings, and prohibits such applications in certain sensitive areas. NVZ also caused a shift in the areas affected toward more productive land use [Burt et al., 1993]. This could have resulted in ploughing of land previously not cultivated, therefore increasing C released from such areas. Moreover, the MacSharry reform was introduced in 1992, which altered the basis of subsidy in the Common Agricultural Policy, started a shift from product support to producer support, and set the introduction of compulsory set-asides [Freibauer et al., 2004]. Similar to the introduction of NVZs, this reform may have caused a shift toward arable production in grasslands capable of sustaining it, and abandonment of arable production in marginal lands, which would maximize subsidies and food production. Finally, the UK faced two major droughts over the periods of 1990–1992 and 1995–1997 [Marsh et al., 2007]. This would result in increased dilution of DOC sources from low-DOC groundwater (which contributes to 63% of the Thames streamflow), and therefore contributing to the increase in DOC concentration. It is notable that mean DOC decreased thereafter. Moreover, the Thames basin was under a severe drought also in 1975–1976 [Marsh et al., 2007], when DOC concentration had another peak.

### 6.3. Making Sense of the 1990s DOC Increase

In the period of 1899 to 1905, when the calibration between DOC and color was performed, the level of urbanization was relatively low, but it increased over all the period. Urbanization increases the amount of sewage effluents discharged to the river, which usually is associated with less colored carbon (given that sewage is typically composed of, among others, carbohydrates and urea, which are nonabsorbing or weakly absorbing DOC substances [Painter and Viney, 1959; Tzortziou et al., 2015]). Therefore, we would expect a likely underestimation of the amount of observed DOC in the river pre-1990, due to the calibration period used. This would result in a greater increase in the observed fluvial DOC over the period when DOC is calibrated from color (1884–1990), as the increasing contribution of less colored carbon with increasing urbanization is not fully accounted for. This would partly reduce the observed steep increase in the observed DOC in the 1990s. Nonetheless, it is worth mentioning that such peaks have been observed before (DOC level during WWII is comparable to the peak in the 1990s) and there is consistency in DOC concentrations over the period when the observed measurements changed from color to DOC. Moreover, the DOC increase in the 1990s could be a result of the implementation of nitrate vulnerable zones (NVZs), which put limits to the application rates of manure and fertilizers, to their timings, and prohibits such applications in certain sensitive areas. NVZ also caused a shift in the areas affected toward more productive land use [Burt et al., 1993]. This could have resulted in ploughing of land previously not cultivated, therefore increasing C released from such areas. Moreover, the MacSharry reform was introduced in 1992, which altered the basis of subsidy in the Common Agricultural Policy, started a shift from product support to producer support, and set the introduction of compulsory set-asides [Freibauer et al., 2004]. Similar to the introduction of NVZs, this reform may have caused a shift toward arable production in grasslands capable of sustaining it, and abandonment of arable production in marginal lands, which would maximize subsidies and food production. Finally, the UK faced two major droughts over the periods of 1990–1992 and 1995–1997 [Marsh et al., 2007]. This would result in decreased dilution of DOC sources from low-DOC groundwater (which contributes to 63% of the Thames streamflow), and therefore contributing to the increase in DOC concentration. It is notable that mean DOC decreased thereafter. Moreover, the Thames basin was under a severe drought also in 1975–1976 [Marsh et al., 2007], when DOC concentration had another peak.

### 6.4. Evaluation of Model Performance

To estimate the predicted DOC, we used export coefficients estimated from national-scale data of the period 2001 to 2007, when the UWWTD had already been transposed into legislation across the UK. The result of this transposition is expected to have improved the wastewater treatment and reduced the input of organic matter into receiving waters. Therefore, the model likely underestimates sewage inputs to the river prior to 2000. Sewage effluents are also a source of more labile DOC [Goldman et al., 2012]. The possibility of the DOC concentration prediction being affected by an enhanced lability of DOC from sewage sources has in effect been accounted for by the manner in which DOC flux is projected and apportioned through the catchment. Equation (6) is based upon the actual observations of over 160 catchments and their behavior, and it includes a loss term which is the average loss term across all sources (including urban) and across a range of different types of catchment.

In spite of the limitations presented above, the model, which was obtained without any optimization or fitting to the Thames data, has an overall mean absolute error (MAE) of 0.57 mg L⁻¹ (Figure 6c), which suggests that these limitations do not impair our conclusions. The overall performance was considered satisfactory, given the long period analyzed. The MAE was chosen, as this objective function is not susceptible to disproportionate influence of few, very high values, as is the case for objective functions based on minimizing the residual sum of squares (e.g., Nash-Sutcliffe efficiency (NSE) or root-mean-square error) [Howden et al., 2011]. The model is able to represent the long-term behavior of DOC, its positive trend, and some of the
major peaks (Figure 6b) which occurred during periods of major land use change, at a large spatial scale such as the Thames basin.

6.5. Limitations of the Study

The focus of this study was to analyze long-term changes in DOC and to capture the dominant modes of behavior of the catchment system. The model considers an annual time step at the catchment scale, constrained by the land use data only being available as an annual summary for the whole catchment. The model is thus able to consider a broad spatial and temporal scale, to give a broad representation of the major sources of C loading contributing to fluvial DOC concentrations and the major processes in a catchment of approximately 10,000 km² over a 120 year period. The variability of DOC increased over the period, partly because the sampling frequency decreased over time. However, here we only considered drivers of the long-term increase in mean DOC concentration, so as to elucidate long-term trends and identify land use drivers. The increase in variability will be the focus for subsequent work.

The transport of DOC from land to river is considerably simplified in the model. We assume that all the C load which reaches the river comes through surface or near-surface runoff. DOC which leaches through deeper soil layers is partly stored in the subsoil and partly reaches local aquifers, but DOC concentration in UK groundwater is found to be generally less than 1.5 mg L⁻¹, and lowest in Chalk aquifers [Stuart and Lapworth, 2016]. Moreover, DOC is nonconservative, as it is consumed by microbiological processes and the travel time through the Thames aquifer has been estimated to be of the order of 30 years [Howden et al., 2011]. Therefore, surface runoff and near-surface runoff are justifiably assumed to be the major pathways of carbon export to the river.

The focus of this study was not deciphering the exact source and composition of DOC in the Thames River, which due to the nature of the data available would not be possible, but understanding the main drivers for its increase and their relative importance. Nonetheless, we expect increased areas of impervious surfaces and artificial storm drainage networks over time to have altered riverine DOC concentrations, due to less contact time between DOC and soil, and therefore, more DOC is exported to the river. During periods of high rainfall, the majority of water from impervious surfaces in urban areas go directly into rivers, without carrying much DOC derived from soil storage [Tian et al., 2012]. Therefore, in periods with high flow, DOC exported to rivers will be less colored. Instead, in periods when the majority of DOC was from terrestrial sources, DOC will be composed mainly by fulvic and humic acids, products of the degradation of lignin and cellulose, which are of darker color and more refractory [Engstrom, 1987].

6.6. Implication of Changes in C Budgets

Over the whole time period analyzed, perturbations in land use and land management were roughly in equilibrium with respect to C release since the 1880s. However, massive conversion of grassland into arable land triggered by the start of WWII caused large releases of C from soils to the river and the atmosphere, reversing decades of carbon sequestration. Policy makers should consider these consequences in future land use decisions, especially in developing countries where the expansion of agricultural land can be more rapid and less regulated [Davis et al., 2015].

Between 63 and 89% of DOC released from the terrestrial biosphere is lost in transit through the catchment over the period of study, which equates to 26.8 kt C yr⁻¹ or a cumulative 3.3 Tg C between 1884 and 2005, part of it to the atmosphere, part of it is buried in in-stream sediments. Our estimate of % C loss is comparable to other estimates for the UK or for all the world. Regnier et al. [2013] have estimated that the total global carbon flux (inorganic and organic carbon) into freshwaters was 2.8 Pg C yr⁻¹, of which 1 Pg C yr⁻¹ reached the tidal limit (i.e., a 64% removal rate). Worrall et al. [2012] measured the net watershed loss of DOM across the UK and found a value of 78%, and it was assumed that this was loss to the atmosphere. Moody et al. [2013], in experimental studies, found a 70% loss of DOC over a 10 day period for UK rivers. Finlay et al. [2016] have summarized net watershed losses of DOC across UK catchments and found a 75% net loss. These losses could potentially be greenhouse gases lost to the atmosphere, which is why in-stream C losses need to be considered in the global carbon assessments for their significant contribution to greenhouse gas emissions [Cole et al., 2007; Battin et al., 2009b; Lapierre et al., 2013; Wit et al., 2015].
7. Conclusions

The major conclusions of this study are as follows:

1. Urbanization has driven the increase of fluvial DOC in the Thames basin for over a century, which is likely linked to the increase in sewage effluent due to increased population.
2. Land use change can release in just a few years carbon stored in soils for centuries, as happened due to the massive grassland ploughing during WWII.
3. The increase in temperature in the Thames basin has not translated into a significant increase in the DOC export to the river.

The increase of DOC in freshwaters ("global browning") is an increasingly worrying phenomenon due to its widespread presence in the northern hemisphere, and for its consequences on aquatic life, leading to increasing costs and complexity of water treatment, and implications for climate change [Oosthoek, 2016]. In the face of a growing world population, this study poses questions on the long-term security of clean water supplies, and on the need to restrict development in sensitive basins, which could release additional carbon to rivers.

Acknowledgments

This work was supported through grants from NERC (NE/J50033/1), the University of Bristol Cabot Institute, and a University of Bristol doctoral studentship award. Thanks to Chris Orton, Department of Geography, Durham University, for drawing the map. The authors are grateful to the Editor, Chris Hopkinson, and one anonymous reviewer for their useful comments. The authors declare no competing financial interests. The DOC data used to generate figures in the manuscript are available at 10.5285/57943561-4587-4eb6-b14c-7adb90dc1dc8.

References

Butman, D. E., H. F. Wilson, R. T. Barnes, M. A. Xenopoulos, and P. A. Raymond (2014), Increased mobilization of aged carbon to rivers by increasing costs and complexity of water treatment, and implications for climate change [Oosthoek, 2016].


