Research Article

Persistent Organic Pollutants in Urban Soils of Central London, England, UK: Measurement and Spatial Modelling of Black Carbon (BC), Petroleum Hydrocarbons (TPH), Polycyclic Aromatic Hydrocarbons (PAH) and Polychlorinated Biphenyls (PCB)

Christopher H. Vane [1, *], Alexander W. Kim [1], Darren Beriro [1], Mark R. Cave [1], Stephen R. Lowe [1, 2], Raquel A. Lopes dos Santos [1], Antonio M.P.J. Ferreira [1], Christopher Collins [2], C. Paul Nathanail [3], Vicky Moss-Hayes [1]

1. British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK; E-Mails: chv@bgs.ac.uk; awk@bgs.ac.uk; darrenb@bgs.ac.uk; mrca@bgs.ac.uk; S.Lowe2@exeter.ac.uk; raquel@bgs.ac.uk; antonio@bgs.ac.uk; vlm@bgs.ac.uk
2. Department Geography and Environmental Science, University of Reading, Reading, RG6 6DW, UK; E-Mail: c.d.collins@reading.ac.uk
3. Land Quality Management, Nottingham NG7 2RD & GHD UK Nottingham, UK; E-Mail: paul@lqm.co.uk

* **Correspondence:** Christopher H. Vane; E-Mail: chv@bgs.ac.uk

Abstract

Total organic carbon (TOC), black carbon (BC), total petroleum hydrocarbons (TPH), polycyclic aromatic hydrocarbons (PAH) and polychlorinated biphenyls (PCB) were determined in 73 surface (0-2 cm) and subsurface (5-20 cm) soil samples taken from a 142 km^2 area in Central London, UK. Soils were assessed to provide a baseline chemistry for site owners, developers, occupiers and regulators involved in understanding the potential risk to human health and the environment. TOC range was 1.75-11.85 % (mean 5.82 %), BC 3.72-

32.71 mg.g^{-1} (mean 13.8 mg.g^{-1}), TPH 72-4673 mg.g^{-1} (mean 443 mg.g^{-1}), Σ^{16}PAH 1.64-421.23 mg.g^{-1} (mean 47.99 mg.g^{-1}) and Σ^{7}PCB 2.56-148.72 µg.kg^{-1} (mean 20.82 µg.kg^{-1}). Surface soils were less polluted than sub-surface soils due to a decline in industry, power generation, coal burning and traffic. PAH and PCB showed a stronger affinity for BC than TOC and were higher than many other international cities. South east London (Greenwich, Woolwich, Deptford) had the highest PAH pollution. Source PAH ratios confirmed a combustion/urban road run-off origin with minor petroleum inputs. Random Forest spatial modelling (machine learning) revealed large scale pollution trends across London soils. Normal background concentrations (NBC) were calculated and compared to risk-based human health generic assessment criteria (GAC). Benz[*a*]anthracene, benzo[*a*]pyrene, benzo[*b*]fluoranthene, and dibenzo[*a*]anthracene exceeded the Land Quality Management GACs for three land uses (residential, allotments and public open space near residential housing). The NBC determined for ∑7PCBs (110 µg.kg^{-1}) and dioxin-like PCB 118 (59 µg.kg^{-1}) exceeded the residential and allotment soil guideline values.

Keywords
Soil; Risk; health; city; urban, contamination; pollution; ecosystem; deprivation; benzo[*a*]pyrene

1. Introduction

The metropolitan area of London has a population of some 9 M people and is one of the world's preeminent cities forming the cultural, economic and governmental hub of the UK. Over 2000 years of human activity have resulted in the accumulation of organic pollutants in soil such as black carbon (BC), petroleum hydrocarbons (TPH) and polycyclic aromatic hydrocarbons (PAH) from coal burning as well as receipt of combustion engine vehicle exhaust particulates. More recently, halogenated persistent organic pollutants (POPs) such as polychlorinated biphenyls (PCB) have been unintentionally released into the environment and transferred into London's surface soils and connected sediments of the River Thames [1-4]. Understanding the spatial and temporal distribution, input-sources and affinities of POPs is important as many are toxic, impact the nervous system or are carcinogenic to humans and wildlife. The presence of POPs in surface and near surface soils potentially constrains the extensive re-development of former industrial sites to housing alongside improvements to London's environment via green-habitat creation, upgraded sewer and urban drainage infrastructure as well as enhanced regulation to improve air quality by 2050 [5].

Petroleum hydrocarbons in urban soils originate from a variety of sources including road run-off, crude and refined oils, refined petroleum as well as dusts and atmospheric particles from factories, incinerators and hydrocarbon fuelled power stations [6-8]. These inputs generate a continuum of simple to complex (monomers-molecular aggregates) that may include aliphatic, monoaromatic, polyaromatic hydrocarbons as well as heteroatom containing resins and higher molecular weight asphaltenes [9]. The precise mix of compound classes entering the environment varies according to oil genesis (basin-type, geological age) degree of refining

(distillation/conversion/finishing) and post environmental release weathering effects such as evaporation and dissolution and microbial alteration [9-11]. Owing to this wide range in molecular composition it is difficult to make assumptions about possible adverse environmental health effects.

Polycyclic aromatic hydrocarbons (PAH) are detected in most urban soils where they occur as a consequence of anthropogenic activities such as road transport, domestic fires, oil refining, coke and asphalt manufacture, coal gas production, non-ferrous metal industries, municipal waste incineration and leaking storage/land-fills [12]. PAHs are associated with a variety of adverse human health effects including suppressing childhood IQ, respiratory health and cancer [6, 13, 14]. In the UK, atmospheric monitoring records for PAH (1970-2003) reveal a switch from residential coal burning, power stations and non-ferrous metal working to road transport (road-run off, vehicle exhaust particulates) [15]. Similarly, isomeric and non-isomeric parent PAH ratios as well as overall concentration patterns confirmed mainly combustion derived PAH (traffic and domestic burning) in near surface soils of $19km^2$ of south east London. Σ^{16} PAH concentrations ranged from 4 to 67 mg.g^{-1} and mean 18 mg.g^{-1} which were broadly similar to some of the more polluted European cities [2]. These results were used to determine normal background concentrations (NBC) of surrogate marker PAH ((e.g. benzo[*a*]pyrene) for comparison against generic assessment criteria (GAC). The GAC providing a risk-based conservative estimate of the surrogate marker PAH concentrations in soil that represent no appreciable or minimal risk to human health assuming chronic exposure for a series of land use and exposure scenarios.

Polychlorinated biphenyls (PCBs) were used in coolants, lubricants, transformer oils, sealants and calking agents from mid 1950s to late 1970s [16]. PCBs were prized for their low reactivity and high chemical and thermal stability. Their properties saw PCBs applied widely in sectors such as industry, transport and commerce, though it was these properties that contribute to their environmental persistence [17]. Peak production was reached in the 1960s before staged phase-out in the late 1970s, following growing concern about their potential for harm and accumulation in the environment, and a global ban through the Stockholm Convention on Persistent Organic Pollutants in 2001[18]. It is in addition to this persistence and accumulative potential that health risks should be considered. PCBs have been found to possess endocrine disruptive qualities associated with developmental disorders in humans and other animals, and some congeners exhibit carcinogenicity. Risk from PCB exposure through multiple pathways can be significant. Soil represents the most significant environmental store of PCBs [16, 19, 20].

Spatial modelling and prediction for the purposes of illustrating the trends in 2D distribution of a contaminant in soil has traditionally been carried out using kriging related techniques [21, 22]. Recently, Hengl et al. [23] summarised the problems associated with kriging and demonstrated the use and advantages of a machine learning method (Random Forest, RF) for spatial modelling where buffer distances from observation points are used as explanatory variables. In other studies Kirkwood et al. [24] used an RF model with high resolution geophysics and remote sensed data as covariates to produce high resolution geochemical maps of South West England and Sekulić et al. [25] successfully demonstrated a similar RF approach using the nearest observations and their distances to the prediction location as covariates. In this study, we have used a RF approach [26] where inverse distance weighted (IDW) covariates are used as the predictor variables.

Although organic pollutants such as PAH, PCB and TPH are frequently measured in UK urban soils to comply with land-use planning system regulation this information is not published in the

peer reviewed literature. In addition, supporting chemical measurements, such as determinations of BC which provide insights into their interaction with natural and co-occurring anthropogenic polymers are rarely undertaken. This study aims to fill these knowledge gaps: (1) provide a baseline geochemical survey of TOC, BC, TPH, PAH and PCB across Central London; (2) compare the concentrations found at two depth intervals and with other major cities; (3) discuss the concentrations with respect to published human health generic assessment criteria (GAC) so this dataset can provide a baseline for long-term environmental monitoring and help to quantify the impact of future pollution changes and/or events; (4) identify, where possible ratios of PAHs and PCBs to describe potential sources of pollution and characterise the likely contribution of emissions; (5) Explore machine learning Random Forest modelling approach to mapping organic pollution trends across Central London.

2. Materials and Methods

2.1 Soil Sampling and Preparation

Urban soils were collected from the Greater London Authority (GLA) encompassing an area of London Boroughs and the City of London, UK (Figure 1). At each site surface soil samples were recovered from four corners and the centre of a 20 m by 20 m square grid at surface 0 to 2 cm (X sample) and subsurface 5 to 20 cm (A sample) using a hand-held Dutch auger. The five auger flights were combined to give ~ 1 kg wet weight which was stored in Kraft wet strength sample bags (125 mm x 250 mm) which in turn were stored in sealable polyethylene outer bags [2, 27]. Upon return to the laboratory, soils were freeze-dried, desegregated and passed through a brass sieve with an aperture of 2 mm before being ball milled (agate balls and lined vessels) to pass a 63 μm sieve. This method was chosen as a statistically robust preparation method for high precision and accuracy PAH quantification [28-30].
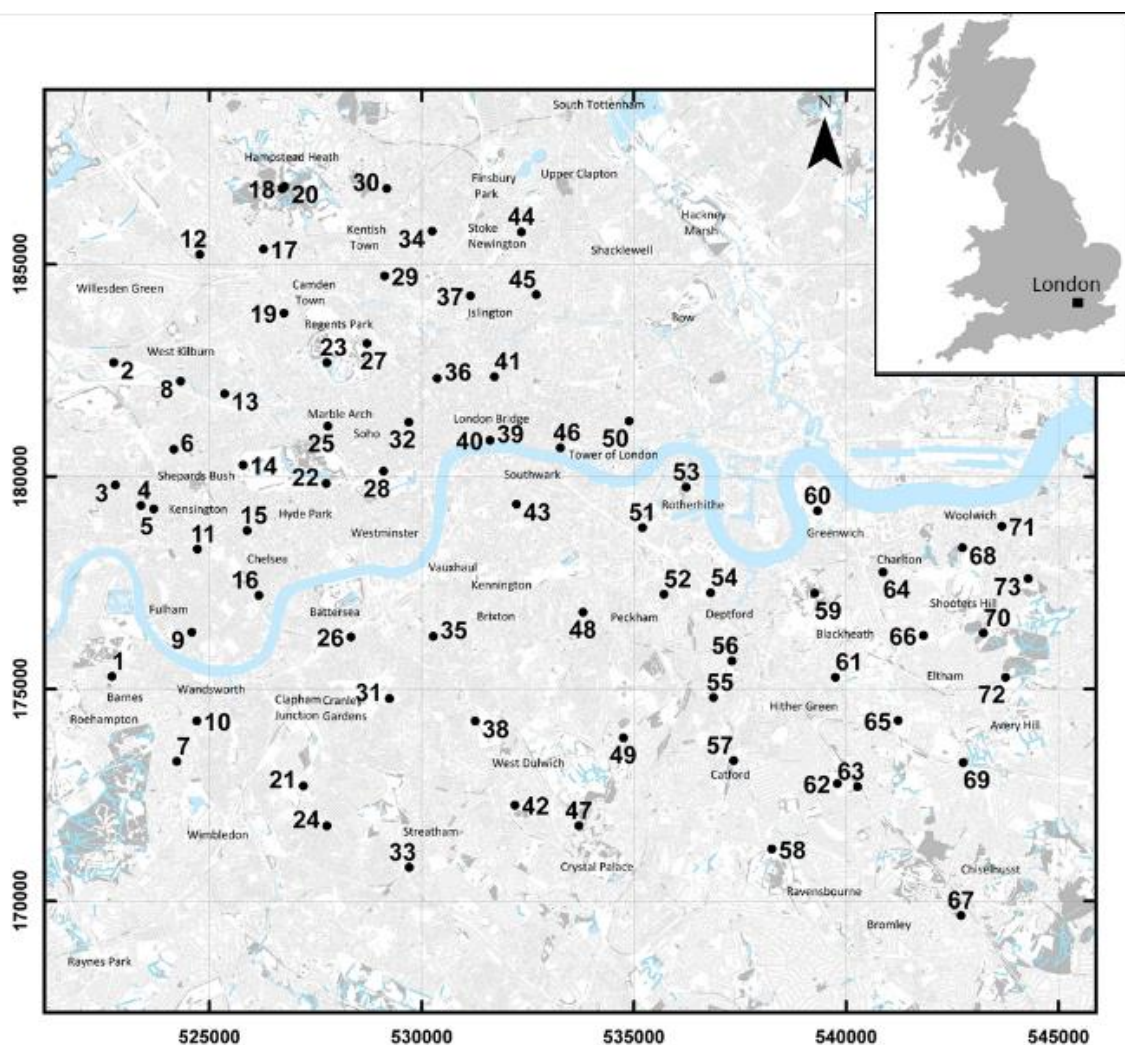
Figure 1 Map showing position of soil sample locations analysed for organic pollutants. Soil data BGS, © UKRI. Contains Ordnance Survey data©Crown Copyright and database rights 2016.

2.2 Organic Carbon, Black Carbon and Total Petroleum Hydrocarbons

Total Organic carbon (TOC %) was determined after acidification with HCl (50% v/v) using a Elementar Vario Max' C/N analyser [31]. The limits of quantification for a typical 300 mg sample were 0.18%. Black carbon was determined by multiple solvent extractions and thermal treatment with elemental analysis of carbon [32, 33]. Total petroleum hydrocarbons were determined after extraction and separation by Iatroscan Mk6 (Supplementary) [7].

2.3 Polycyclic Aromatic Hydrocarbons

Polyaromatic hydrocarbons (PAH) (10 g) were spiked with deuterated standards, extracted with dichloromethane/acetone (1:1 v/v) using ASE, reduced in volume and isolated for analysis by solid phase extraction (SPE) using an identical scheme and operating conditions as those reported for urban soils and sediments from Clyde Basin and New York [7, 8]. Polyaromatic hydrocarbons (PAH) were measured using a Varian 3800 gas chromatograph (GC) coupled to a Varian 1200L triple

quadropole mass spectrometer operating in full scan mode (ionization energy 70 eV, mass range 47-500 amu). Application (1 μL) was by splitless injection from 0-0.7 min and thereafter a split 1:20 was applied (injector temperature 280 °C). The GC was fitted with a fused silica Varian VF-35MS column (30 m length × 0.25 mm i.d. × 0.1 μm film thickness). The GC oven-temperature programme was 60°C (1 min. isothermal) to 320°C (at 6°C.min^{-1}.) and held isothermally at 320°C for 10 min and helium was used as carrier gas (1 mL.min^{-1}). Quality control was achieved by duplicate analyses of a single certified reference material (NIST 1941b) at a minimum of every 10 samples. The LOQ was determined peak height at 5 times signal / background ratio.

2.4 Polychlorinated Biphenyls

Soils (10 g) were spiked with 50 μL of PCB 34 (18.92 ng.μL^{-1}), PCB 62 (19.77 ng.μL^{-1}), PCB 119 (18.92 ng.μL^{-1}), PCB 131 (19.24 ng.μL^{-1}) and PCB 173 (18.81 ng.μL^{-1}), and 50 μL of a recovery standard solution of PCB 19 (20.01 ng.μL^{-1}) and PCB 147 (19.31 ng.μL^{-1}). After 1 h equilibration, spiked soils were combined with clean sand (1:2) and copper powder added before ASE extraction using acetone/hexane (1:1 v/v). Extracts were reduced to 3-4 mL and mixed with H_2SO_4 to release humic bound PCBs [28, 34]. The supernatant was retained and reduced to 0.25 mL before transfer to a pre-filled Na_2SO_4 SPE cartridge (Agilent, Bond Elute TPH, 500 mg). The first fraction was eluted with 1.5 mL of pentane and discarded. The second fraction was eluted with 6 mL of hexane propanol (97:3 v/v) and spiked with 100 μL of the analytical standard solution (PCB 29, 9.47 ng.μL^{-1}) and PCB 157 (9.53 ng.μL^{-1}) and reduced to 200 μL under a stream of nitrogen. Analysis was by combined gas chromatography-mass spectrometry (GC-MS) using a Fisons GC8000 gas chromatograph coupled to a Fisons MD800 single quadrupole mass spectrometer in full scan mode (ionisation energy 70eV, mass range 39-600 amu)[28, 34, 35]. Quantification of PCBs was achieved using selected ions for the ICES 7 PCB (International Council of the Sea: congeners 28, 52, 101, 118, 138, 153 and 180) which represent a large proportion of most commercial formulations, span a wide range of chlorination (3-7) and are widely considered a key suite enabling comparison with other soil and sediment studies.

2.5 Spatial Modelling Interpolation

The data analysis was carried out using the R programming Language [36] and associated libraries. The "sf " R library [37] was used to attribute the deprivation data to the sampling points and prediction grid. The "ranger" R library [38] was used to carry out the RF modelling. The Boruta R library [39] was used to select the significant predictor variables in the RF model. Preliminary checks (not reported) show that after the top 5 to 10 most important IDW predictor combinations, inclusion of further IDW predictors does not meaningfully improve the root mean square error of prediction (RMSEP). Details of the four stages of modelling and interpolation are presented in the supplementary.

3. Results & Discussion

3.1 Organic Carbon

London surface soils TOC (0-2 cm) ranged from 3.05-11.85 %, mean of 6.58 % and median of 6.07 %, standard deviation 2.11 % and counterpart subsurface soil TOC (5-20 cm) ranged from

1.75-9.61 % mean of 5.06 % and median of 4.95 %, standard deviation 1.66 % (Figure 2, Table 1). Taking both soil depths together, TOC ranged from 1.75-11.85 %, mean of 5.82 % and median TOC of 5.42 % (n=146) (Table 1). Evaluation of the TOC showed that surface soils contained higher amounts of organic carbon than sub-surface soils (t-test; P=<0.001) (Table S1). Variation in TOC is probably due to the diverse range of current and historical land-use and inclusion of two depth horizons at each site, which, not-withstanding physical disturbance common to most urban soils, probably differ in age and extent of decomposition of the more labile organic compounds and polymers [40, 41]. Conversely, factors such as the incorporation of geological parent material (pre-Anthropocene TOC) were discounted since the entire study area is underlain by London Clay (Eocene) and organic rich peats (Holocene) only outcrop at a few locations along the River Thames foreshore and associated historical flood plain [42, 43]. The higher TOC in the top soil as compared to the sub soil is probably best explained by enhanced incorporation of living and decayed organic matter (e.g. fine leaf litter, hair-roots, soil humus) rather than an increase in anthropogenic carbon. A similar near surface, rise in TOC ascribed to greater amounts of fine roots and decomposed plant remains has been previously observed in upper horizons of urban soils and sediments from Staten Island, New York, USA and vegetated mud islands on the River Thames, London, UK [3, 8, 11].

Table 1 Summary of Total Organic Carbon (TOC), Black Carbon (BC), Total Petroleum Hydrocarbons (TPH), Polyaromatic Hydrocarbons (PAH), Polychlorinated Biphenyl (PCB) concentrations in Soils of London, England, UK.

Description	TOC (%)	BC (mg.g^{-1})	TPH (mg.kg^{-1})	PAH (Σ^{16}) (mg.kg^{-1})	PAH (Σ^{27}) (mg.kg^{-1})	PCB (Σ^{7}) (µg.kg^{-1})
All depths (n=146)						
Mean	5.82	NA	443	35.20	47.99	
Median	5.42	NA	323	21.23	29.37	
Min-max	1.75-11.85	NA	72-4673	1.64-421.23	2.42-421.23	
Std dev.	2.04	NA	491	43.52	55.93	
Soil 0-2 cm (n=73**)**						
Mean	6.58	13.80	421	29.51	39.29	*19.35
Median	6.07	12.40	298	19.00	25.64	*7.53
Min-max	3.05-11.85	3.72-32.71	72-4673	1.64-184.32	2.42-235.27	*2.56-148.72
Std dev.	2.11	6.77	582	35.00	45.13	*28.24
Soil 5-20 cm (n=73)						
Mean	5.06	NA	466	40.89	54.69	**21.77
Median	4.95	NA	346	31.92	42.07	**10.40
Min-max	1.75-9.61	NA	89-2618	3.64-336.10	5.22-421.23	**2.56-135.81
Std dev.	1.66	NA	382	50.24	64.36	**28.46

NA BC not analysed for this depth interval; PCBs analysed on sub-set of the 73 sites *n=39 sites, **n=61 site
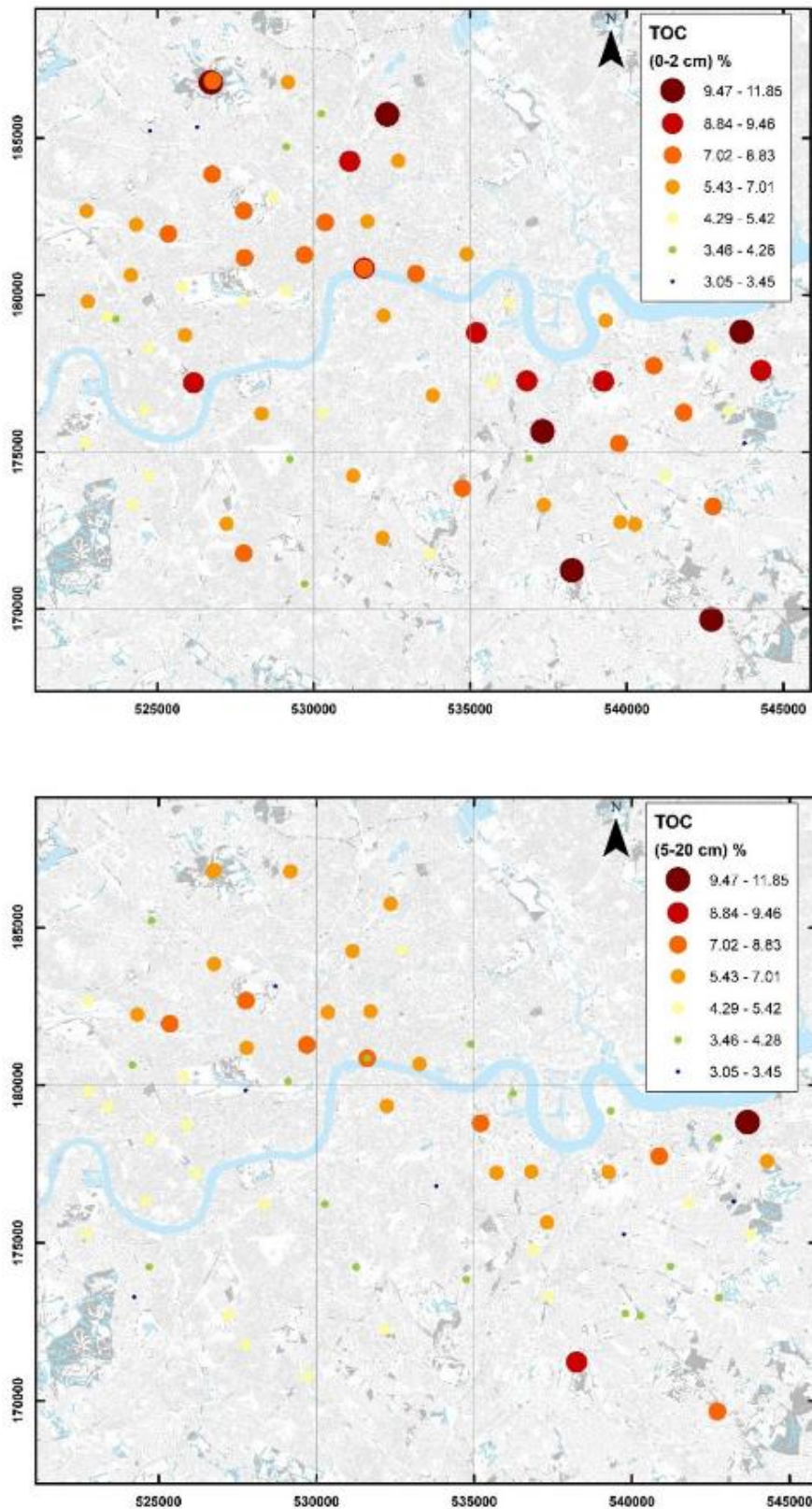
Figure 2 Map of London soil total organic carbon content (TOC %). Soil data BGS, © UKRI. Contains Ordnance Survey data©Crown Copyright and database rights 2016.

Other studies of UK urban soils also report considerable variation in organic carbon content, for example sub soils (5-20 cm) from south east London (n=73) report TOC ranging from 3.04 to

21.77 %, mean 7.19 %, median 6.63 % and eighty four soils from across Glasgow city spanned 2.79 % to 17.97 % with a mean of 6.15 % and a median concentration of 5.34 % [2, 27]. Therefore, the variable and sometimes high TOC % presented here for London is commensurate with other urban studies (Figure 2). Both lower and higher soil organic carbon content have been reported for the smaller, formerly industrialised urban centres of Stoke (n=737, range 0.4 to 42.5 %, mean 7.3, median 6.8 %) and Coventry (n=808, range 0.9 to 18.6, mean 3.5, median 3.3). However, some caution in comparison of TOC different studies is required. This is because the TOC % were obtained using Loss on ignition (LOI) at 450°C, a cost effective, but less accurate method for organic carbon that is influenced by a range of confounding factors including: 1) loss of structural water and oxyhydroxides in clays ; 2) incomplete carbon combustion at low temperatures (~450 °C) and loss of inorganic carbonates (>440°C) [44, 45]. Whilst these limitations are not necessarily significant in sand dominated soils they are relevant to any soil containing clay mineral particles such as London soils or those with high inorganic carbon content (e.g. carbonate).

Due to its importance in sustaining primary agricultural and horticultural productivity as well as underpinning other ecosystem services, soil organic carbon (SOC % synonymous with TOC %) has been measured and evaluated in non-urban rural UK soils [32, 46, 47]. The 2007 summary data for top-soils (0 to 15 cm) from the Countryside Survey reported mean English soil organic carbon (from LOI) of 7.65 % [48]. Lower TOC % were reported for arable and horticultural soils (3.8 %), whereas slightly higher TOC were found in improved and neutral grassland soils (6.77 and 8.36 %), acid grasslands (25.79 %), shrub/heathland (26.46 %) and broad-leaf/mixed woodland (13.01 %) as well as coniferous woodland 19.94 % [48]. Comparison of TOC values suggest the London soils have higher TOC than UK agricultural and horticultural soils, a similar TOC to improved and neutral grasslands and lower TOC compared to semi-natural acid grasslands, broad leaf and coniferous woodlands, shrub/heathland and wetlands. This mid-table position in the overall UK soil organic carbon ranking contrasts with studies from Germany (Rostock), USA (New York and Baltimore), Moscow (Russia) which reported elevated values in urban land-use domains compared to their rural equivalents [49-52]. The higher urban organic carbon content was attributed to a range of factors including; 1) addition of admixtures during landscaping; 2) residential settlement history and composting; 3) inhibition of microbial litter decay due to an increase anthropogenic polyaromatic hydrocarbon concentrations; and 4) higher use of wood/coal fuels linked to latitude and climate and temperature. However, other studies such as those comparing sites New York, USA and natural and sealed urban sites in Seoul, South Korea found 66 and 27 % less TOC in the urban domain [53, 54]. Therefore, the data presented here for London does not fully support either extreme notion (higher/lower than rural counterparts) but rather suggest that urban soils are part of a continuum of UK soil carbon inventory.
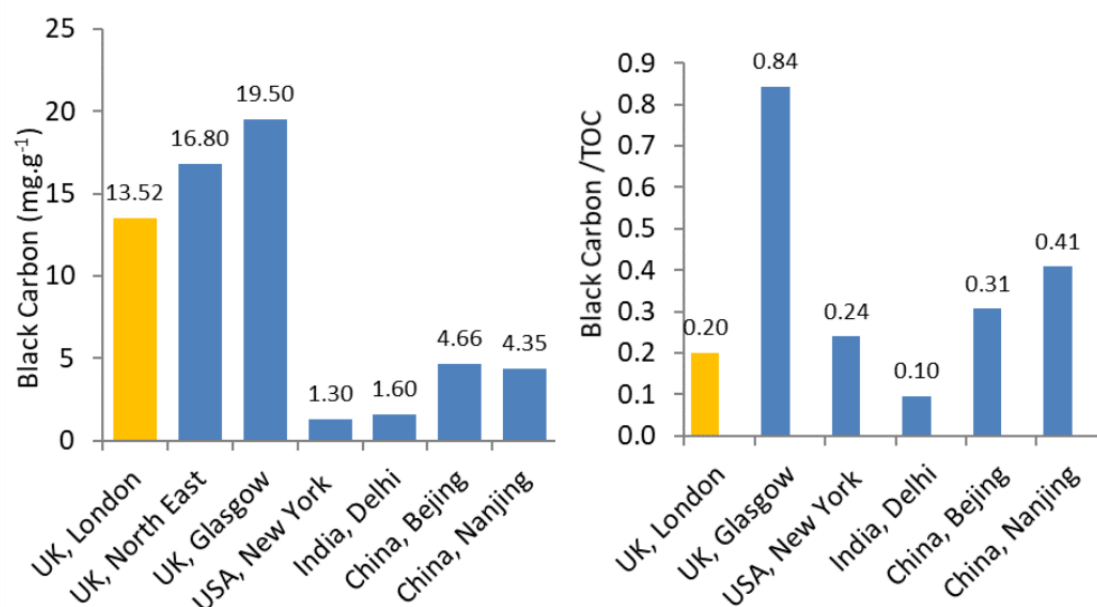
3.2 Black Carbon

Black Carbon (BC) surface soil concentrations are shown in Figure 3 and Table 1. The BC ranged from 3.72 to 32.71 mg.g^{-1}, mean 13.80 mg.g^{-1} and median of 12.40 mg.g^{-1} with a positive linear correlation between BC and TOC (R^2 0.45) indicating that BC only partly controls London urban soil TOC. One plausible explanation for this association being that soil TOC is dominated by natural organic matter such as fine plant roots, decaying organic matter (leaf litter, microbial tissues) and repolymerised decay products and residues (humus) [55]. In contrast, Urban soil BC most likely

originates from anthropogenic activities such as burning of wood or coal or pyrolysis of wood (charcoal) as well as other contributors such as vehicle exhaust particulates [32, 56]. Black carbon to total organic carbon ratio (BC/TOC) ranged from 0.07 to 0.4 with a mean of 0.2 (Figure 4). In general, BC/TOC ratios of <0.1 are taken to indicate incorporation of burnt biomass whereas higher BC/TOC ratios of 0.4 are attributed to input from combustion of fossil fuels. London surface soil (0 to 2 cm depth) BC/TOC ratios fall between these criteria suggesting incorporation of both combustion fossil fuel and burnt biomass sourced organic matter or alternatively some unknown industrial or household waste source inputs.
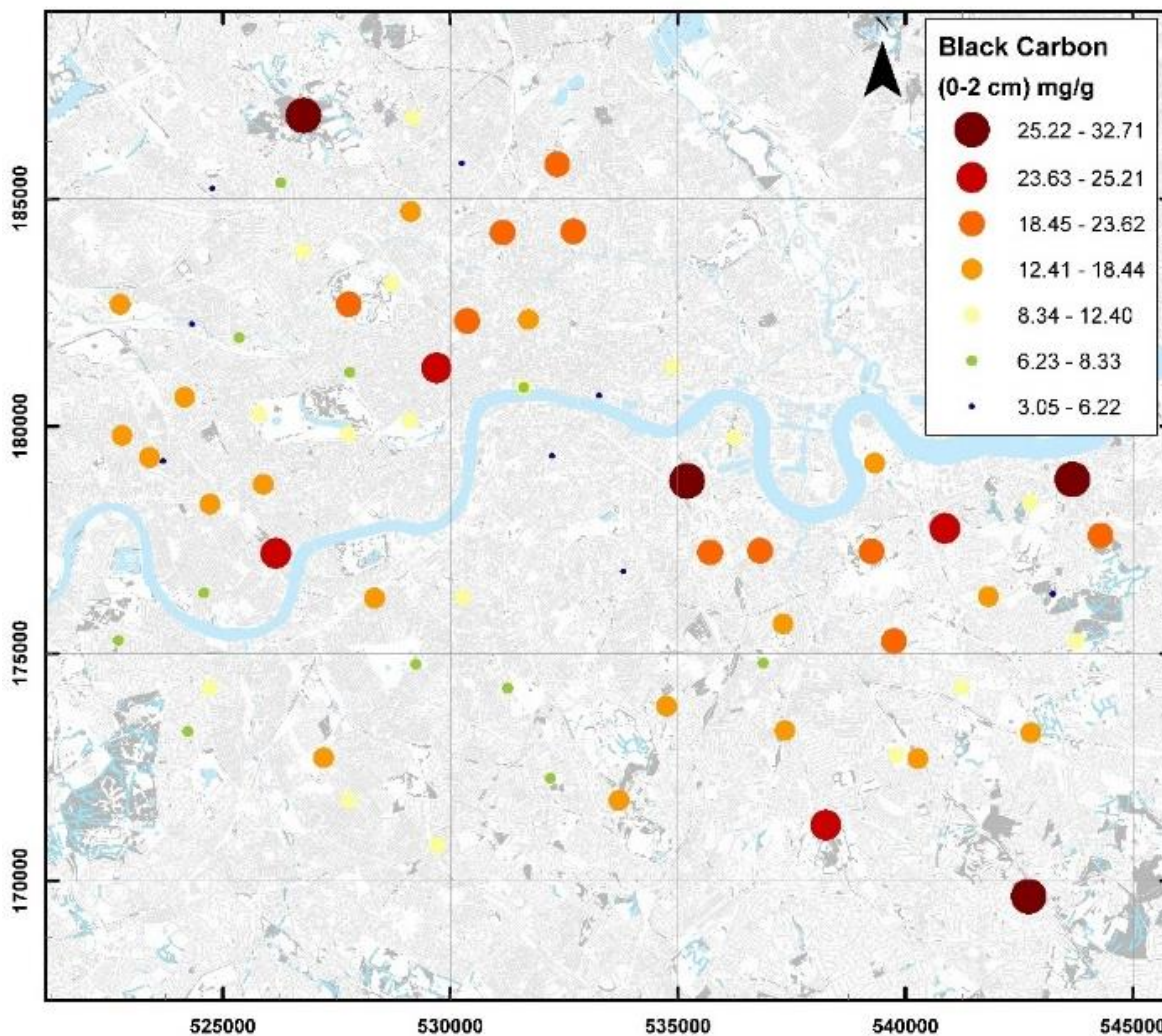
Figure 3 Concentration of black carbon (BC) in surface soils of London. Soil data BGS, © UKRI. Contains Ordnance Survey data©Crown Copyright and database rights 2016.

Figure 4 Comparison of black carbon in urban soils [32, 56-61].

London soil BC concentrations are similar to those reported for other UK urban-industrial centres, namely, Glasgow, Coventry and Stoke (4.6 to 17.7 mg.g^{-1}) and about ten times higher than woodland and grassland background soils of the UK (n=31) which varied from 0.45 to 3.13 mg.g^{-1} (mean 1.36 mg.g^{-1}) (Figure 4) [32, 62]. Conversely a regional scale study of 55 urban soils in North East England reported even higher concentrations of 16.8 mg.g^{-1} in the topsoil decreasing to a median of 5.7 mg g^{-1} at 1 m depth [56]. It should however be borne in mind that although lower than NW UK, a comparison of BC concentrations from other world cities suggests that London soil BC content is relatively high, possibly reflecting protracted BC accumulation from legacy industries, long history of domestic coal burning (until 1960s) combined with the high density of coal and gas fired power stations situated along the London reaches of the river Thames; mainly 1850s to 1980s [1, 63]. For example, a study of Beijing urban soils from 0 to 20 cm BC ranged from 0.37-12.7 mg.g^{-1}, mean 4.66 mg.g^{-1} and BC/TOC ratio ranging from 0.04-0.91 with a mean of 0.31 (Figure 3). Overall, the high BC concentrations, low BC/TOC value and weak positive correspondence to TOC % supports the notion of multiple industrial and domestic BC inputs rather than a single dominant source (e.g. wood/charcoal burning only).

The occurrence of BC at high concentrations in surface soils has both positive and negative environmental implications. Owing to its high porosity BC and particularly the charcoal fraction can adsorb pollutants including PAH, PCB, dioxins and polybrominated diphenyl ethers (PBDE) as well as heavy metals thereby reducing their bioavailability to ecology and humans [58, 62, 64-66]. Similarly, the presence of BC in soils has also been reported to increase cation exchange capacity as well as physically stabilise humic acids which in turn increases soil fertility [67]. However, the more condensed soot fraction can be detrimental to animal and human health. For example, inhalation of particulate organic matter containing nanoparticles of BC adversely effects lung function and has been linked to cardiovascular mortality and clinical effects including ventricular,

repolarization, changes in blood pressure heart rate variability, endothelial dysfunction. Epidemiological studies have identified that distance of the home to the nearest main road is important and a recent evaluation of airway macrophage carbon in London's commuting cyclists showed that inhaled dose of BC was linked to time taken to travel and route (traffic density)[68]. Whilst it isn't possible to directly link soil hosted BC concentration to atmospheric particulates and effect on human health it is reasonable to assume that the variations in BC accumulating in soil are at least in part linked to atmospheric sources (e.g. soot, vehicle exhaust, road dust) although other anthropogenic inputs (coal, ash, slag) cannot be entirely ruled out because the BC methodology employed cannot discriminate BC fractions as it encompasses a broad range of condensed aromatic structures.

3.3 Total Petroleum Hydrocarbons (TPH)

Concentrations of petroleum hydrocarbons (TPH) in London top soil (0 to 2 cm, *n*=73) ranged from 72-4673 mg.kg^{-1} (std deviation 582 mg.kg^{-1}), mean 421 mg.kg^{-1}, median 298 mg.kg^{-1} with saturated hydrocarbon (aliphatic) concentrations ranging from 70-3576 mg.g^{-1}, mean 277 mg.g^{-1}, median 162 mg.g^{-1} and aromatic hydrocarbon concentrations from < 3-1097 mg.g^{-1}, mean of 144 mg.g^{-1} median of 108 mg.g^{-1} (Figure 5). The biogenic lipid/resin concentration in the top soils ranged from 114-13153 mg.g^{-1}, mean 991 mg.g^{-1} and median of 771 mg.g^{-1}. Counterpart, London sub soil (5 to 20 cm, *n*=73) TPH ranged from 89-2618 mg.g^{-1}, mean 466 mg.g^{-1}, median 346 mg.g^{-1} with saturate hydrocarbons ranging from 89-1,118 mg.g^{-1}, mean 283 mg.g^{-1}, median 214 mg.g^{-1} and aromatic hydrocarbon concentrations from 17-1500 mg.g^{-1}, mean of 194 mg.g^{-1} median of 139 mg.g^{-1} and the biogenic lipid/resin concentrations ranged from 95 to 4,776 mg.g^{-1}, mean 753 mg.g^{-1}, median 585 mg.g^{-1} (Figure 5). When both soil horizons are considered together London soils (*n*=146) concentrations of TPH varied from 72-4,673 mg.g^{-1}, mean of 443 mg.g^{-1} and median of 323 mg.g^{-1}. Assessment of soil TPH data confirmed that that the deeper sub soils (5-20 cm) contained greater amounts of TPH than surface soils (0-2 cm) (t test; P=<0.002) which substantiates a relative decrease in London's anthropogenic organic and heavy metal pollutants since its peak in 1940s-1980's [1, 3].
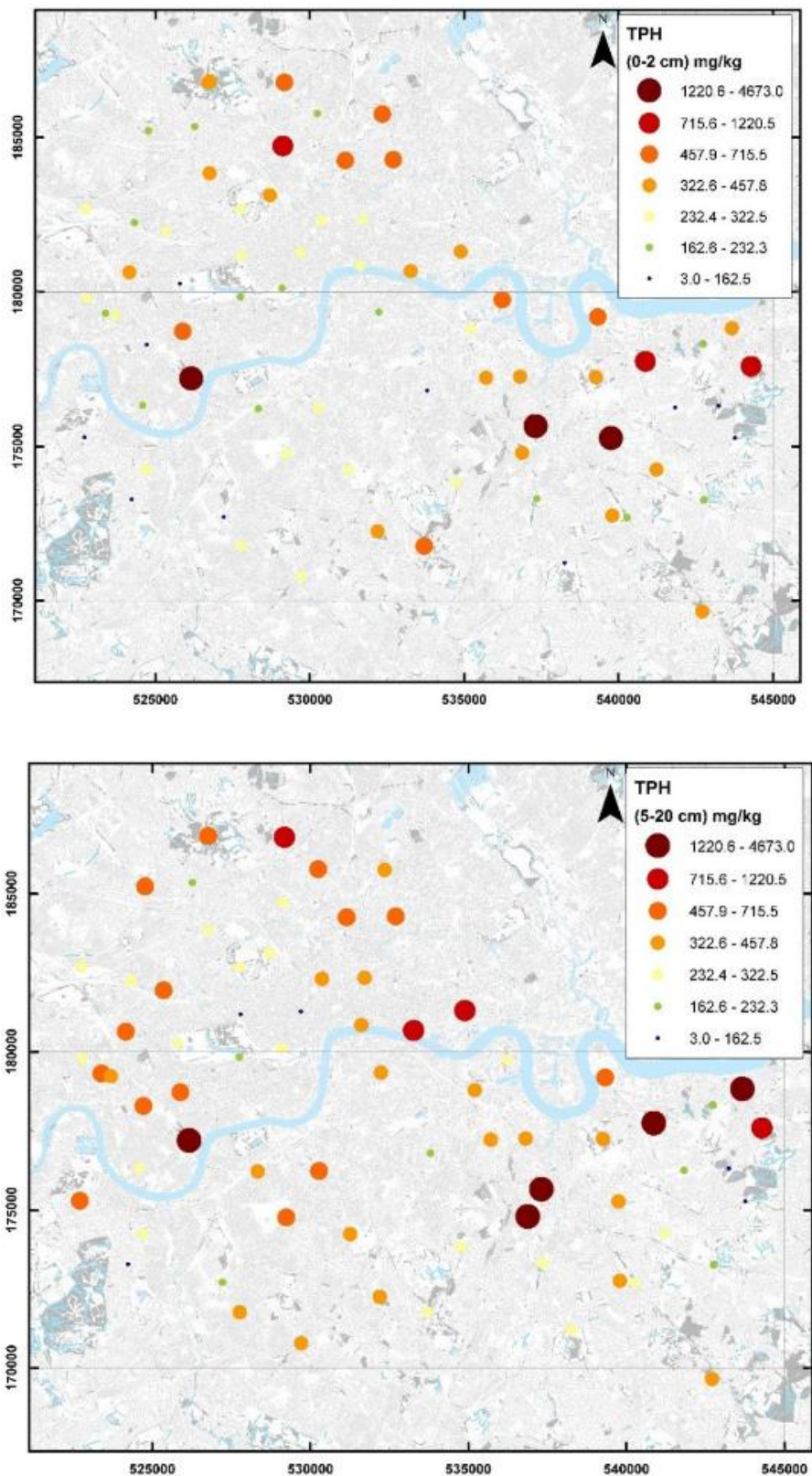
Figure 5 Total petroleum hydrocarbons (TPH) in London soils. Soil data BGS, © UKRI. Contains Ordnance Survey data©Crown Copyright and database rights 2016.

Before proceeding with an evaluation of London soil TPH concentrations it should be borne in mind that soils considered largely devoid of pollution generally contain a background of

hydrocarbon compounds owing to their biosynthesis by most organisms and subsequent incorporation of partially degraded biomass in soils and sediments (e.g. grass stems, tree roots, epicuticular leaf waxes, cell walls (bacteria, moulds, fungi), algae and animal tissues) [9, 69-71]. For example, unpolluted soils and sediment intervals from mangrove soils of south China and marsh soils of Virginia, New Jersey and New York, USA showed TPH concentrations ranging from 100-300 mg.g^{-1} [8, 35, 72]. Similarly, unpolluted river sediments from the upper rural reaches of River Clyde, UK found TPH spanning 3-260 mg.g^{-1} with a mean of 68 mg.g^{-1} and unpolluted sediment River Derwent (Australia) ranged from 10-<500 mg.g^{-1} [7, 73, 74]. In New England, USA, the state agencies defined as 373 mg.g^{-1} whereas for regulatory, non-risk based evaluations in Australia and New Zealand TPH guidelines of <280 mg.g^{-1} (no concern) and 550 mg.g^{-1} (of concern) are used within a weight of evidence assessment for coastal sediments [73, 75]. Conversely, soils receiving appreciable diffuse and or point source (e.g. crude oil, refined oils and fuels, industrial discharge) hydrocarbon pollution typically contain TPH >1000 mg kg^{-1} [76]. For example, sediments from polluted urban tributaries of Glasgow, Scotland, UK yielded a mean TPH of 2,779 mg.g^{-1} including concentrations as high as 37,879 mg.kg^{-1} and oil spill contaminated marsh soils from Staten Island, New York, USA report TPH 2,500-9,586 mg.kg^{-1} [7, 8]. Further, a more nuanced understanding of urban soil TPH concentrations has been gained from industrial waste and fresh end-member materials analysed in parallel with urban soils [27]. The TPH concentrations of the source end members were highly variable and ranked as follows: creosote impregnated railway sleeper (23,150 mg.kg^{-1})>coal (9,259 mg.kg^{-1})>bitumen road dust>(884 mg kg^{-1})>metal processing industrial slag>(181 mg.kg^{-1})[27]. Therefore, in this current work a TPH concentration >500 mg.g^{-1} was taken to indicate the presence of anthropogenic sourced TPH.

Benchmarking against the TPH >500 mg.g^{-1} anthropogenic hydrocarbon criterion revealed exceedances of 11 surface soils (0-2 cm) and 20 sub soils (5 to 20 cm) indicating the presence of appreciable anthropogenic saturate and aromatic hydrocarbons. The most likely cause being incorporation of wastes from fuel spills, legacy coal particles, road dust bitumen, diffuse traffic exhaust or other unknown legacy domestic/municipal or industrial wastes. A recent survey of 84 urban sub soils (5-20 cm) collected across the city of Glasgow reported lower ranges of 75-2,505 mg.g^{-1} and a means 388 mg.g^{-1} as compared to those here for London soils (e.g. mean 421 and 426 mg.g^{-1})[27] (Figure 6). In view of the inherent inhomogeneity of urban soils, variable end-member concentrations and different histories of urbanization and legacy industries of London and Glasgow the ranges and means of ~400 mg.g^{-1} are remarkably consistent and provide a broad benchmark by which future improvements in soil quality could be benchmarked against.
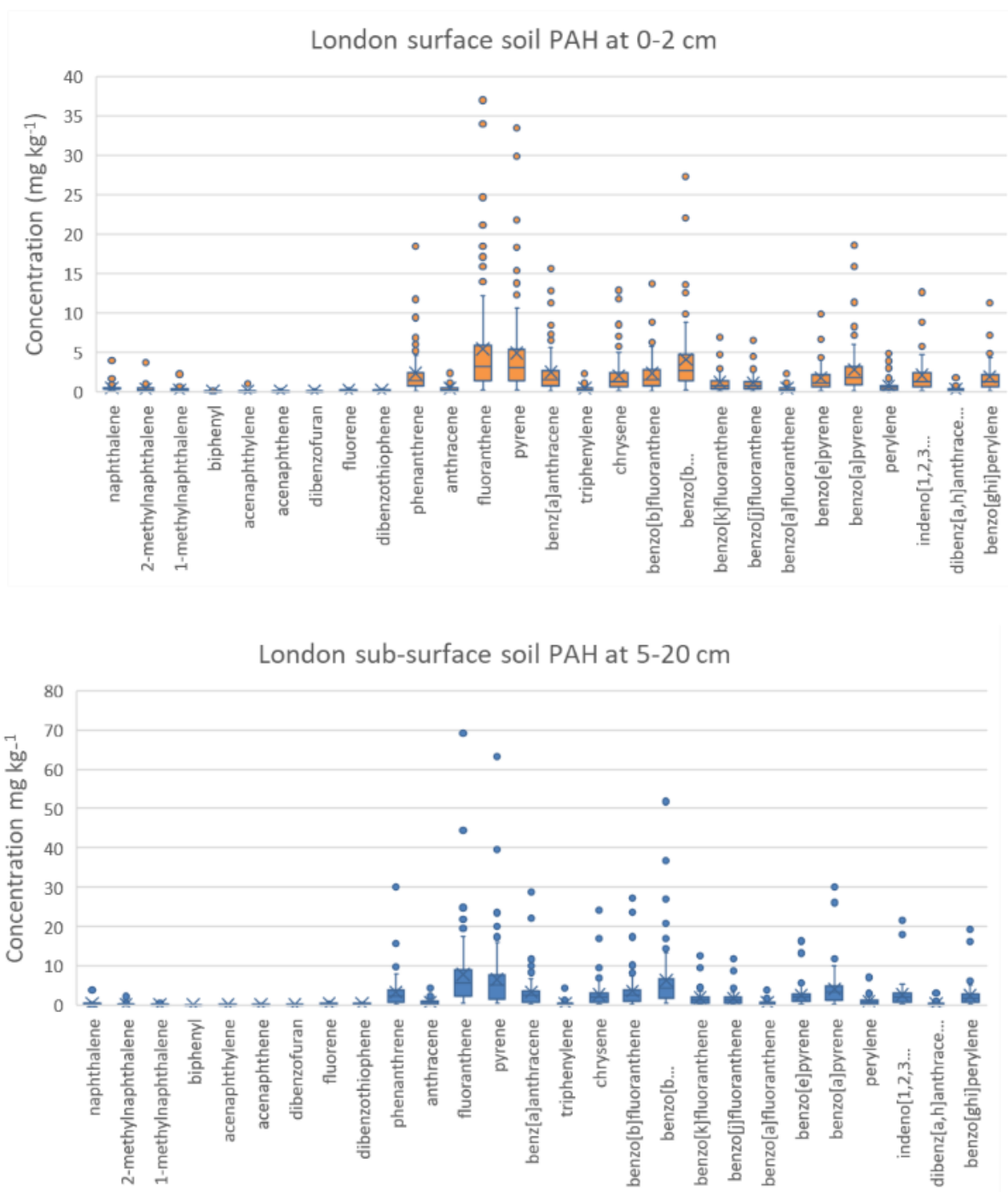
Figure 6 Box and whisper plot of PAH concentrations for London soils; displaying outliers (circle markers), highest and lowest non-outliers (upper and lower whisker limits), upper and lower quartiles (upper and lower box limits) and median values (horizontal red line).
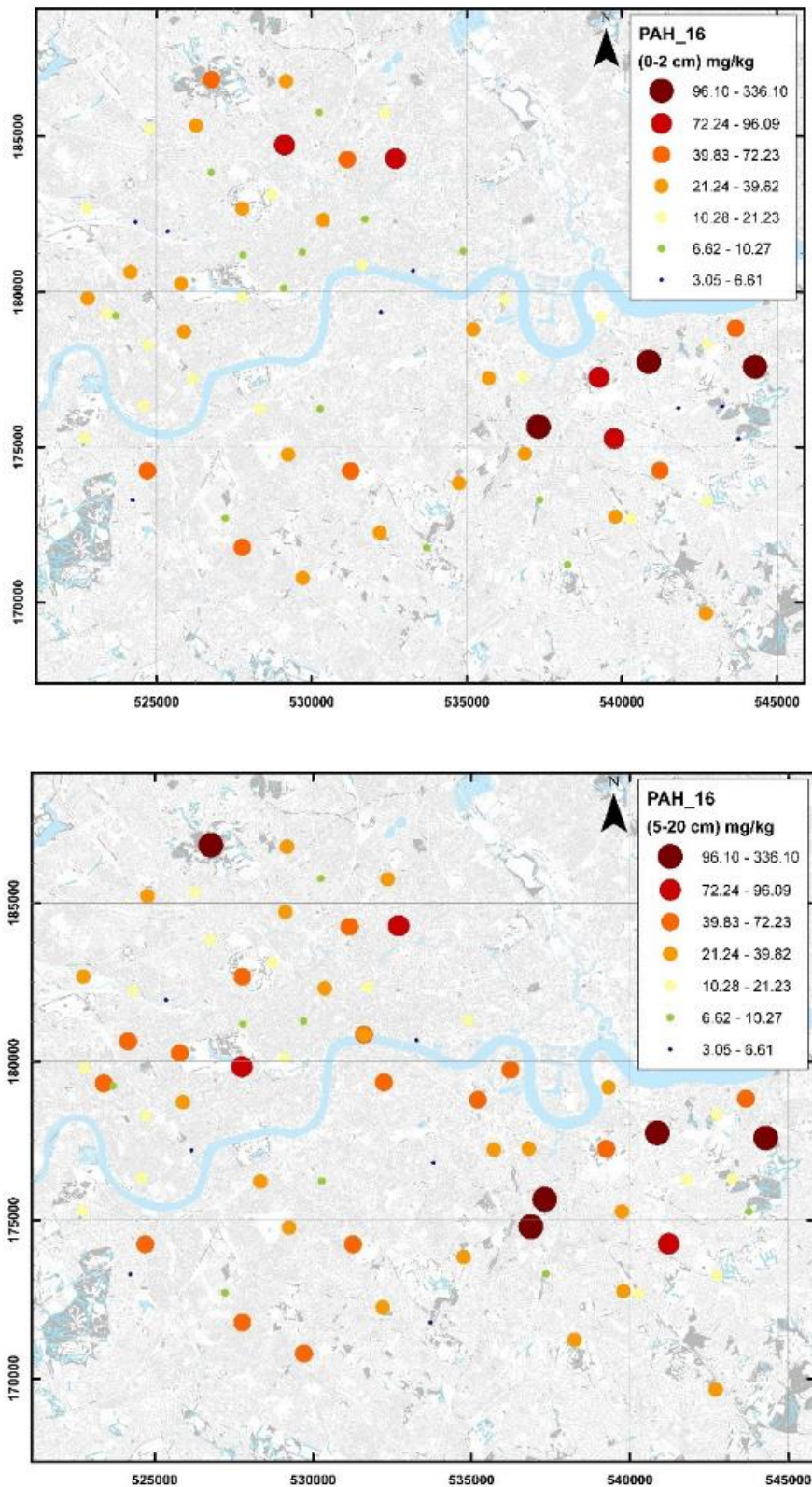
3.4 Polyaromatic Hydrocarbons (PAH)

Concentrations of Σ^{16}PAH in London top soil (0 to 2 cm) ranged from 1.64 to 184.32 mg.g^{-1} with a mean 29.51 mg.g^{-1}, median 19.00 mg.g^{-1} and $\Sigma27$ PAH ranged from 2.42 to 235.67 mg.g^{-1} with a mean 39.29 mg.g^{-1} and median 25.64 mg.g^{-1} (Table 1). The Σ^{16} PAH in London sub soil (5 to 20 cm)

varied from 3.64 to 336.10, mean 40.89 mg.g^{-1}, median 31.92 mg.g^{-1} and Σ27 PAH concentrations ranged from 5.22 to 421.23 mg.g^{-1} with a mean 54.69 mg.g^{-1} and median 42.07 mg.g^{-1} (Table 1, Figure 6). When taken together the London soils (n=146) had a mean Σ^{16} PAH of 35.20 mg.g^{-1} and Σ^{27} PAH of 47.99 mg.g^{-1} (Table 1). The Σ^{16}PAH, Σ^{27}PAH and benzo[a]pyrene (BaP) concentrations were significantly higher in the sub soil as compared to the surface soil (t-test, p=<0.001) (Table S1).

The highest soil PAH concentrations were observed in the south east boroughs of Greenwich, Woolwich, Deptford which is in keeping with the areas ongoing road traffic congestion issues and or industrial legacy (Figure 7) [1]. In addition, sites spanning central London's Hackney, Stoke Newington and Islington had elevated PAH concentrations as compared to Camden, Hampstead, Kensington, Fulham, Marylebone, Paddington, Wandsworth which have to a lesser or greater extent residential land-use histories. The relative contributions made by 27 individual PAH in London top and sub-soils are presented in boxplots (mean denoted by cross, median value by a line, together with min-max values and outliers using the 1.5 times interquartile range criterion (Figure 6). These show minor contributions of about 0.5 mg.g^{-1} of the total PAH mass of low molecular weight PAH (2–3 ring) such as naphthalene, acenaphthylene, 1-methylnaphthalene and major contributions of about 3 to 5 mg.g^{-1} of the total PAH mass from higher molecular weight PAH (4–6 ring) dominated by fluroanthene, pyrene, benzo[a]pyrene, benzo[b]fluronanthene, benzo[b/k]fluronanthene, benzo[e]pyrene and benzo[a]pyrene as well as benzo[ghi]perylene. However, it should be borne in mind that not all 4 and 5 ring PAH were dominant, for example triphenylene, perylene and dibenz[a,h]anthracene were at about 1 mg.g^{-1} (Figure 6).

Figure 7 Concentration of polyaromatic hydrocarbons (Σ16 USEPA PAH) in London soils. Soil data BGS, © UKRI. Contains Ordnance Survey data©Crown Copyright and database rights 2016.

Although sub soil (5-20 cm depth) total Σ^{27}PAH concentrations were higher than top soils the relative contributions across the molecular weight/ring sizes did not change. This similarity in

relative contribution across 2-6 ring PAH supports the notion of a fairly consistent urban background source(s) input and/or possibly reflects rapid near surface PAH weathering effects (e.g. leaching, photolysis, volatilisation, microbial mediated decay) and or mixing. Although PAH concentrations in soils generally decrease with depth, some studies have reported elevated PAH at depth due to sequential deposition of different contaminated industrial wastes yielding rise-peak and fall profiles similar to that observed in water deposited sediment cores or the addition of burnt building material and rubble from the second world war (1939-1945) [8, 12, 77]. In the light that London was heavily bombed during WWII and that the majority of sites have had multiple land-uses both explanations are plausible. Overall the PAH concentration data confirms a general recent decline in PAH pollution in surface soils (0-2 cm) as compared to deeper older counterpart (5-20 cm).

On a national basis (UK), the mean concentrations and ranges of PAH for central London rank above those for other studies. For example, 76 soils from 5-20 cm of east London (Thamesmead-Erith-Abbey Wood) had a Σ^{16}PAH ranging from 4 to 67 mg.kg^{-1}, mean of 18.6 mg.kg^{-1} and median of 14.0 mg.kg^{-1} and 84 soils from Glasgow, Scotland showed a mean of 32.4 mg.kg^{-1} and median of 12.5 mg.kg^{-1} [2, 27]. The UK Environment Agency Soil and Herbage Survey of 2007 reported $\Sigma22$ PAH for 42 English urban soils. These ranged from 0.79 to 55.1 mg.kg^{-1} with a mean 25.7 55 mg.kg^{-1}, median 10.5 mg.kg^{-1} and a standard deviation of 83.9 mg.kg^{-1} whereas English rural soils (n=182) $\Sigma22$ PAH ranged from 0.04 to 16.8 mg kg^{-1}, mean 0.86 mg.kg^{-1}, median 0.93 mg.kg^{-1} and a standard deviation of 2.74 mg.kg^{-1} [78]. Therefore, both soil depth intervals (0-2 and 5-20 cm) presented here for London (means 29.51 and 40.89 mg.kg^{-1}) are slightly higher than previous studies of metropolitan areas and about 35 times higher than rural soils of England. The most plausible explanation for the higher concentration of PAH in central London soils as compared to other UK urban areas is that London has by far the greatest population of any UK city (~10 M) and therefore in all likelihood the greatest PAH input from domestic wood and coal combustion and traffic. Also, the entire city received appreciable historical aerial input from the numerous coal and oil fired power stations that flank the Thames in London (e.g. Battersea, Bankside (Tate Modern), Barking, Fulham, Greenwich, Woolwich, Chelsea) as well as direct waste fill from fuel and industrial processing plants (gas works/coke production/aluminium production/tar-asphalt production). Therefore, the elevated PAH concentrations are entirely in keeping with this legacy.

On an international basis the PAH concentrations reported here are higher than most comparably sized metropolitan / urban centres. For example, surface soils (0-2 cm) from Manhattan, New York (*n*=27) yield Σ^{16}PAH content ranging from 1.89 to 22.91 mg.kg^{-1}, mean 6.63 mg.kg^{-1} and median 6.53 mg.kg^{-1} whereas counterpart sub soils (>15 cm) spanned 0.94 to 34.05 mg.g^{-1} mean 5.10 mg.kg^{-1} median 4.67 mg.kg^{-1} [59]. In Beijing, China, soils from urbanised districts (*n*=233, 0 to 10 cm) ranged from 0.09 to 131.41 mg.kg^{-1} mean 1.22 mg.kg^{-1} and median of 0.68 with a standard deviation of 1.80 mg.kg^{-1} and surface soils (0-5 cm) from Delhi, India impacted by traffic sourced report Σ^{16}PAH 1.0 to 9.65 mg.kg^{-1} mean of 4.69 mg.kg^{-1} [79, 80]. In contrast, higher concentrations have been recently reported from Ruhr, Germany (Σ^{16} PAH 0.1 to 70 mg.kg^{-1}, mean 44.6 mg.kg^{-1}, median 31.1 mg.kg^{-1}) where the high PAH content was directly attributed to the presence of visually identifiable anthropogenic components including coal, charcoal, ash, rubble and industrial tailings [81]. Therefore, the similar PAH content here in the London soils broadly supports the notion inferred from BC and TPH concentrations that the majority of soils have received a mixture of contemporary and historical anthropogenic inputs.

The relationship between soil TOC and total PAH content was evaluated using a Pearson correlation coefficient. Surface (0 to 2) and subsurface (5 to 20 cm) soils showed a weak positive correlation of 0.25 and 0.29 and BC to Σ^{16}PAH of 0.34. The moderate associations are probably due to a variety of factors including: 1) wide range TOC inputs in urban soils; 2) plant biomass (grass, leaves, stem, root bark) dominated by structural biopolymers (cellulose, lignin, tannin, suberin) that are devoid of PAH; 3) the genesis and transport of PAH (soot, coal fragments/dusts, particulate matter) which may hinder or assist surface sorption, diffusion, entrapment to the mineral fraction; 4) variability of BC form which in turn imparts for PAH [29, 82-84]. These findings are consistent with other studies of urban soils/sediments suggest that PAH concentration normalisation to TOC or SOM may in some cases be inappropriate due to the variable importance of individual processes controlling pollutant-mineral exchangeability and divergent natural organic matter[85]. Overall, the weak positive correlation of BC to PAH suggests that the BC in-itself is not a dominant source of PAH.

3.5 Source Apportionment of Polycyclic Aromatic Hydrocarbons (PAH)

Allocation of soil PAH to pyrogenic, petrogenic or biogenic source categories is commonly achieved using a combination of descriptors such as: 1) distribution pattern of parent and alkylated PAH; 2) ratio 2-3 to 4-6 ring parent PAH; 3) presence of specific PAH that indicates a specific source (e.g. biological perylene) and; 4) isomeric and non-isomeric PAH ratios [12, 86-88]. The first approach is based upon the premise that PAH formed slowly at low temperatures yield substantially higher amounts alkylated than parent PAH (e.g. crude oil, coal) whereas rapid high temperature formation yields higher amounts of parent as compared to alkyl-substituted PAH (e.g. Power station and industrial emissions/traffic exhaust/soot). Similarly, the fourth source apportionment term arises from the principal that one isomer has greater thermal stability and can therefore indicate high temperature combustion (pyrolytic synthesis) as compared to low-temperature petrogenic (unburnt fossil fuels) input. However, before proceeding it should also be noted that most environmental forensic studies recognise considerable intra and inter-source ratio variability which can alongside weathering effects confound unequivocal allocation [7, 89].

The distribution pattern of parent to alkyl-substituted PAH can be used to elucidate anthropogenic PAH inputs in soils and sediments. Inspection of the mean surface soil parent (C0-C3 or C0-C2) for naphthalene, fluorene, phenanthrene, fluoranthene + pyrene and chrysene reveal a complex pattern which is partly sloped suggesting pyrogenic sources but also exhibits a bell-shaped distribution profile for phenanthrene P0-P2 series possibly indicating some petrogenic (unburnt fossil fuel) influence (Figure 8). The reverse slope of the naphthalene N0-N2 series is unusual and maybe explained in part by the semi-volatility of N0 (Figure 8). In contrast the sub-soil (5-20 cm) interval shows phenanthrene, fluoranthene + pyrene and chrysene homologous series are sloped suggesting mainly pyrogenic source(s) (Figure 8). Measurement of PAH in German urban soils mixed in the laboratory with bituminous coal (petrogenic source) and a tar inputs (pyrogenic source) reported that a third V-shaped distribution pattern is observed in fluoranthene + pyrene (C0-C4) between 40:60 to 94:5 % and that this distribution pattern is also found for summed ΣC0 and ΣC2 alkylated PAH with maximal decrease occurring for C1 PAH [81]. In this current study, inspection of the London soil parent-

alkylated distribution profiles reveals a somewhat similar V-shaped pattern for chrysene series at both soil depth intervals and a possible V-shaped profile for naphthalene in the sub soil confirming multiple anthropogenic sources in both (Figure 8). The ratio of parent 2-3 to 4-6 ring PAH provides a complementary means to parent-alkyl distribution pattern for source identification since unburnt fossil fuels (petrogenic origin) yield 4-6 ring /ΣPAH of <0.40 whereas burnt (pyrogenic) inputs 4-6 ring /ΣPAH of >0.50 [90]. In this current study surface soil 4-6/ΣPAH ranged from 0.77 to 0.92 (mean 0.87) and sub soil ranged from 0.61-0.99 (mean 0.88) which suggests that all the soils were dominated by PAH formed from pyrogenic processes such as fuel combustion, vehicle exhaust, burning of wood (Figure 8).
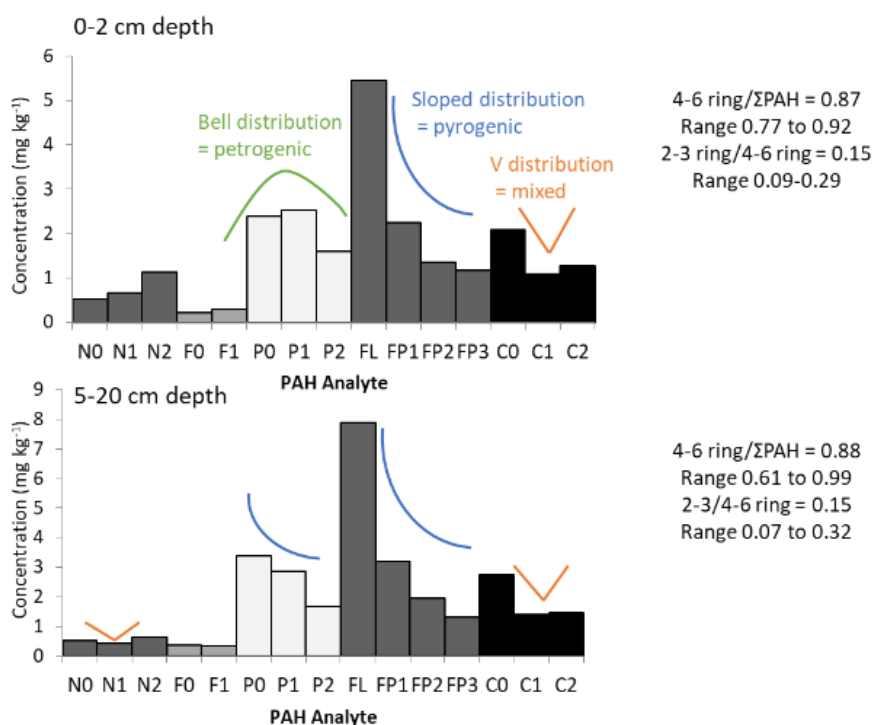

Figure 8 Mean distribution of parent and alkylated polycyclic aromatic hydrocarbons (PAH) in London soils (see Figure 1). N=naphthalene, F=Fluroenes, P=phenanthrenes, FL=fluroanthenes, FP=combined fluroanthenes+pyrenes, C=chrysenes. Degree of alkylation is indicated by number.

Particulates from vehicle traffic exhaust can yield high concentrations of benzo[*ghi*]perylene relative to benzo[a]pyrene such that values >3.14 are used to infer traffic sourced PAH in urban soils and sediments [77, 78]. In this current study surface soils (0-2 cm) benzo[*ghi*]perylene to benzo[a]pyrene ratios ranged from 0.58 to 1.24, mean 0.72, median 0.68 and sub-soil (5-20 cm) ranged from 0.28 to 1.19, mean 0.62, median 0.59. Therefore, none of the soils exceeded the boundary value of 3.14 which on face value could be taken to indicate low amounts of traffic sourced PAH. However, this interpretation seems rather implausible given that the majority of soil sample sites were situated in close proximity to roads including a few from some of London's busiest intersections (e.g. Marble Arch (25), London Bridge, A100 (40), A200 motorway). This therefore calls into question ability of benzo[*ghi*]perylene to benzo[*a*]pyrene ratio to identify traffic sources over and above a complex urban background. Concern over the dependability of

the ratio is further supported by inspection of sites away from major roads (e.g. Hampstead Heath (sites 18, 20) with values of 0.61 and 0.53 and 0.73 and 0.70 that were not significantly lower than equivalents from outside the park (sites 12, 17, 30) (Figure 1). Furthermore, inspection of sites with the highest ratios (>1) revealed some to be situated close to roads (site 48 Peckham and site 46 Tower of London) whilst others were collected close to Railway lines (e.g. site 16, Stamford Bridge football ground (Chelsea FC) (Figure 1). The ratios benzo[*ghi*]perylene to benzo[*a*]pyrene presented herein are consistent with an earlier study of sub soils from south east London (range 0.6 to 1.1, mean of 0.8) and when considered together suggest the ratio is unable to indicate enhanced traffic input over and above the urban background urban. This lack of fidelity may arise for a number of reasons including; 1) the denominator benzo[a]pyrene is varying according to a different source input; 2) alteration of the original source input ratio caused by post depositional weathering effects; 3) the ratio only works well when there are few other PAH inputs.

Isomeric ratios are commonly applied to infer the major PAH source inputs in urban soils and associated urban waterway sediments [2, 8, 27, 28, 59, 77, 88]. The bi-plots presented in Figure 9 utilise fluorene to pyrene ratio (x axis) to suggest petrogenic (uncombusted fossil fuel) as compared to pyrolytic (combusted) sources whereas benzo[a]anthracene to chrysene, benzo[a]anthracene to benzo[a]pyrene and chrysene to benzo[a]pyrene (y axis) are all to a lesser or greater extent used to indicate extent of distillation processing and or combustion or mixing [88]. London surface soil bi-plots show a tight cluster of fluorene to pyrene ratios which maybe taken to indicate a combustion of petroleum (e.g. traffic exhaust) and coal and vegetation which are typical of UK urban background sources. The one exception to this tight grouping and input assignment pertains to soil 71 from Woolwich-Plumbstead a site situated on the boundary of the Royal Arsenal Woolwich, a former military complex that included explosive testing, chemical laboratories as well as armaments manufacture and storage (~1800 to 1994) (Figure 1). Given the unusual land-use chemical association it's entirely plausible that the atypical PAH ratio occurs as a result of combusted input from one or more of the former facilities. Inspection of the London sub-soil (5-20 cm) bi-plots shows a slightly broader spread of ratios spanning petroleum and vegetation/coal combustion sources and comparing with contributions from carburetted coal water gas (CWG), coal tar and urban background (e.g. road-run off). The greater range of ratios in the sub soil as compared to the surface soil is most likely due to both the broader range of past industrial activities and also the greater time span represented by the deeper interval (e.g. 5-20 cm to 0-2 cm).
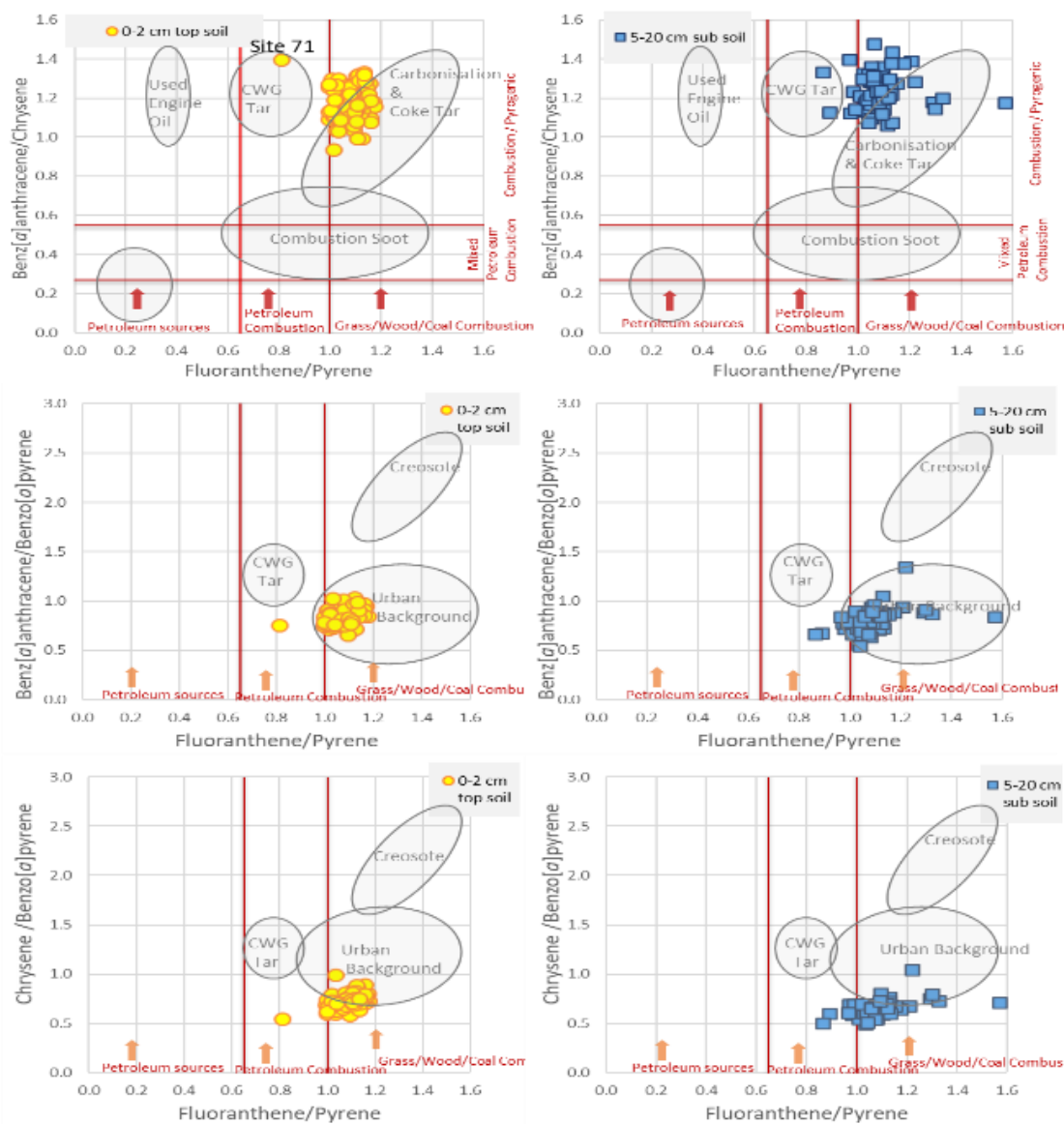
Figure 9 Indicative source assignment of London soil polycyclic aromatic hydrocarbons (PAH) based on common diagnostic ratios based upon thermodynamic stabilities and anthropogenic processes [8, 27, 88, 91, 92].

3.6 Polychlorinated Biphenyls

Concentrations of Σ^7 PCB in London top soil (0-2 cm, *n*=39) ranged from 2.56-148.72 µg.kg⁻¹ (std deviation 28.24 µg.kg⁻¹), mean 19.35 µg.kg⁻¹, median 7.53 µg.kg⁻¹ (Table1, Figure 10). London sub soil (5-20 cm, *n*=61) PCB concentrations ranged from 2.56-135.81 µg.kg⁻¹, mean 21.27 µg.kg⁻¹, median 10.40 µg.kg⁻¹. When both soil intervals are considered together London soils (*n*=100) Σ^7PCB varied from 2.56-148 µg.kg⁻¹, mean of 20.82 µg.kg⁻¹ and median of 9.68 µg.kg⁻¹. Comparison of data confirmed that that the deeper sub soils (5-20 cm) contained greater amounts of PCB than surface soils (0-2 cm) (t-test; P=<0.001) an attribute most likely driven by improved UK regulation between 1972 to 1977 before enclosed uses were highly restricted from 1981 onwards (Table S1). The decrease in PCB concentration in the surface soils parallels the decline in PCB in the near surface portions of sediment cores from undisturbed river island (Chiswick Ait) situated in the River Thames, London [3]. However, it should be borne in mind that although statistically

significant the difference in London Σ^7PCB between top and sub soil isn't as pronounced as might be expected given half-lives of PCB congeners and peak use in early to mid-1970s which in turn suggests either enhanced protection, bioturbation/physical mixing from lower intervals and or continued low level atmospheric input (e.g. PCB contaminated dusts).
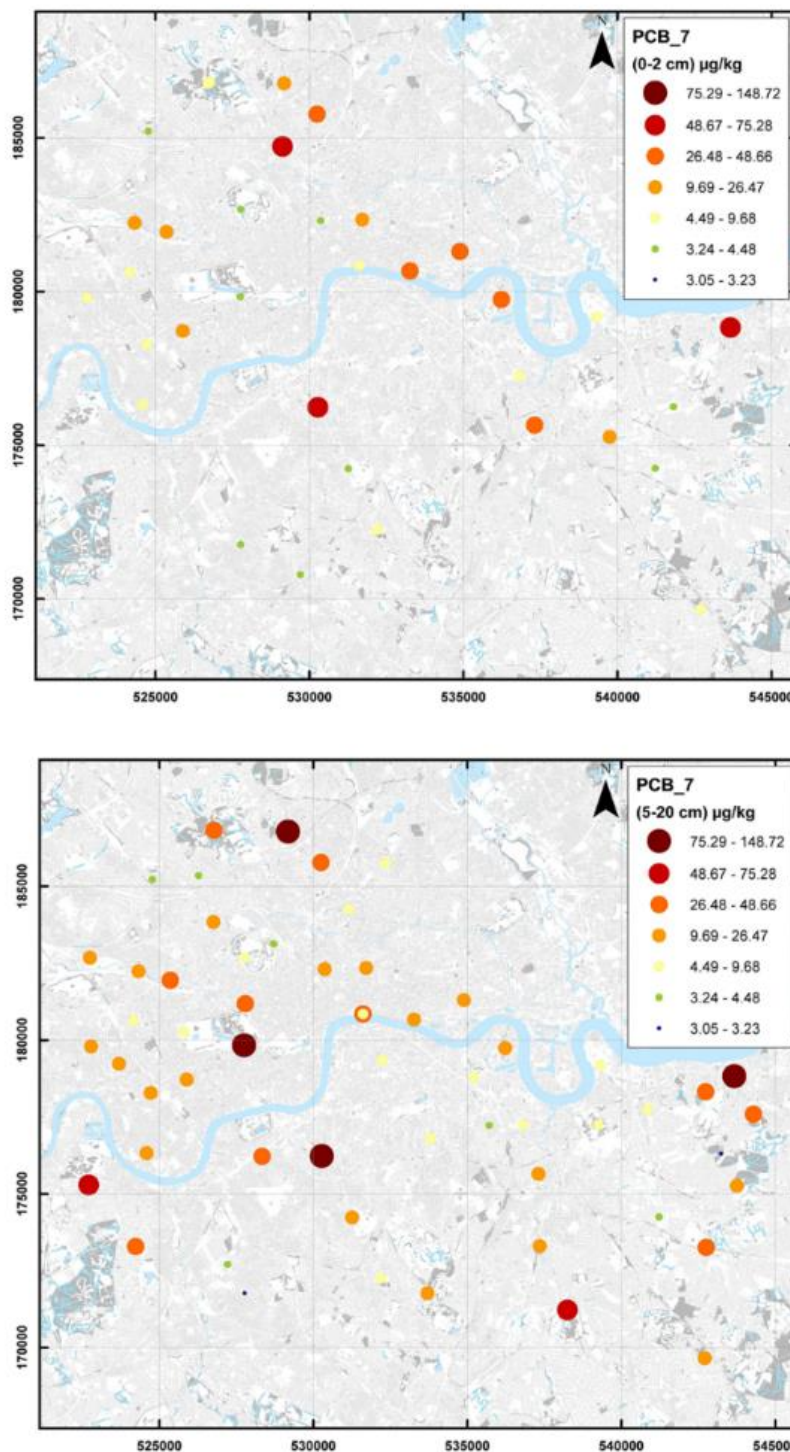
Figure 10 Concentration of polychlorinated biphenyls (Σ^7 PCB) in London soils. Soil data BGS, © UKRI. Contains Ordnance Survey data©Crown Copyright and database rights 2016.

On a UK basis, the ranges, mean and median concentrations of PCB for central London soils are a little narrower than data presented from surveys of urban soils which do not target any particular industrial site or land-use (e.g. urban background). For example, previous studies by Organic Geochemistry (BGS) reported south east London soils (*n*=76) contained Σ^7PCB ranging from 1-750 µg.kg^{-1}, mean of 21 µg.kg^{-1}, median 4.9 µg.kg^{-1} and Glasgow soils (*n*=84) Σ^7PCB content was reported to vary from 2.2-1052 µg.kg^{-1} with a mean 52 µg kg^{-1} [3, 27]. The central London PCB concentrations shown herein are higher than those presented from the UK Environment Agency Soil and Herbage Survey of 2007 (n=42) which stated Σ^7PCB ranging from 0.52-30 µg.kg^{-1}, mean 3.19 µg.kg^{-1}, median 1.18 µg.kg^{-1} for urban soils and median values of 1.29 µg.kg^{-1} for industrial soils and rural soils 0.66 µg.kg^{-1}.

Comparison of median Σ^7PCB concentrations reveals that the central London soils of this current study are four (surface 0-2 cm) and five (sub-surface 5-20 cm) times higher than reported for urban English soils and approximately eleven and sixteen times higher than PCB content of English rural soils. This clearly indicates that London soils contain an elevated background of PCB compared to other urban UK soils; presumably due to receipt of multiple industrial emissions (power generation, waste incineration, paper manufacture, oil refining, metal works (non-ferrous), cement works, chemical manufacture and textile production) and possibly secondary release from demolition dusts and concrete sealants [93]. Former landfill sites and associated soils are known to transfer high amounts of POPs including PCBs to adjacent waterways and soils, nevertheless London soils are slightly higher than those reported from a former landfill site soils, Brighton, England which varied from 1.7-13.2 µg.kg^{-1} [94]. In contrast, the London soils are far lower than those UK landfill near Cheshire, England which reported Σ^7PCB 67-215 µg.kg^{-1} (capped or partially capped) and concentrations as high as 208-140,000 µg.kg^{-1} [95].

Background PCB concentrations in London soils presented in Figure 10 are broadly similar to those reported in other European urban soils such as those from Torino, Italy Σ19 14 µg.kg^{-1} (range 1.8-172 µg.kg^{-1}), Bayreuth, Germany, Σ^{12}PCB 13 µg.kg^{-1} but are higher than studies of urban and peri-urban soils in Rouen, France 0.21 µg.kg^{-1}, Aveiro, Portugal Σ^{19} 7.9 µg.kg^{-1} (0.62-73 µg kg^{-1}), Uppsala, Sweden Σ^{19}PCB 5.7 µg.kg^{-1} (2.3-77 µg.kg^{-1}) Ljubljana, Slovenia Σ^{19}PCB 6.8 µg.kg^{-1} (2.8-48 µg.kg^{-1}) and Moscow, Russia Σ^7PCB 10.5 µg kg^{-1} (3.1-18.5) [96-98]. Conversely, 36 soils and road dusts from 12 sites in Madrid, Spain contained higher PCB concentrations (mean) Σ^7PCB 113.4 µg.kg^{-1} (Σ^{15}PCB 100-1000 µg.kg^{-1}) than those encountered herein. Inspection of Asian urban soil PCB concentrations reveals higher and lower values due to different regulation histories, enhanced volatilisation of lower chlorinated congeners due to higher temperatures and wide variety of PCB containing municipal and manufactured wastes. For example, the London PCB concentrations were far higher than those of Korba District, Chhattisgarh, India, Σ^{28}PCB mean 9.21 µg.kg^{-1} median 0.65 µg.kg^{-1} (3.25-25.2 µg.kg^{-1}) and Ulsan and Gwangyang, Republic of Korea Σ^7PCB 0.38 and 0.59 µg kg^{-1} but were far lower than those reported for 25 provinces and cities in Chinese with mean PCB of 4984 µg.kg^{-1} (0.3-123, 467 µg.kg^{-1}) although it should be noted the latter data-set was strongly skewed by a few high values from points sources in Zhejiang province [99-101]. Therefore, on an international basis, London soil PCB concentrations are higher than most European urban soils but lower than highly polluted Chinese soils.

The relationship between soil TOC and Σ^7PCB was evaluated using a Pearson correlation coefficient; surface and subsurface soils showed very weak positive correlations (r=0.01; r=0.12) whereas a stronger correlation (r=0.35) was found for Σ^7PCB and BC. Combined this suggests that

PCBs have little affinity for natural TOC whereas the anthropogenic BC fraction, which was about a factor of 5 lower than TOC, has a moderate positive association to PCB. The latter association is most likely explained by sorbtion of PCB onto BC surfaces and/or physical entrapment in charcoal porosity [62, 64]. However, another more general explanation may be that urban soils are subject to a wide variety of pollutants from common sources (co-emission) which in turn yields a positive association although doesn't necessarily infer interaction. This result differs from previous evaluation in rural soils from remote rural and woodland areas of UK and Norway which reported a stronger association of PCB and co-planar PCB with TOC as compared to BC [62]. This difference in PCB association (TOC vs. BC) is most likely due to the very low BC content (factor of 100 to 300 lower BC than TOC) in the earlier study, which suggest that BC plays a more important role in partioning and sorption of PCBs in urban soils than in rural soils.

A summary of each of the 7 measured PCB congener concentrations for both depth intervals are presented in Figure 11. The mean congener rank order for surface soils was PCB no. 138 > 118 > 28 >153>101>180>52 whereas subsurface soils PCB congeners occurred in the order 101>28 >118>138>153>52>180. The switch in PCB congener ranking most likely reflects changing source input across years (0-2 cm) as compared to decades (5-20cm) as well as photodegradation, variable adsorption to mineral matrices and aerobic as well as microbial degradation/alteration. In addition, a range of other factors including the "grass hopper" effect where lower molecular weight congeners such as PCB 28 are re-emitted by volatilization from soil to air to soil maybe important modifiers of the PCB distribution pattern. The central London PCB rankings presented herein differ from those from the UK Environment Agency Soil and Herbage Survey of 2007 which reported urban English towns/cities in the order 153>138>180>118>101>52>28 as compared to 153>138>180>118>28>101>52 in rural soils. One plausible explanation for this difference maybe that the earlier study examined 0-5 cm depth interval which differs from the 0-2 and 5-20 cm interval evaluated here.
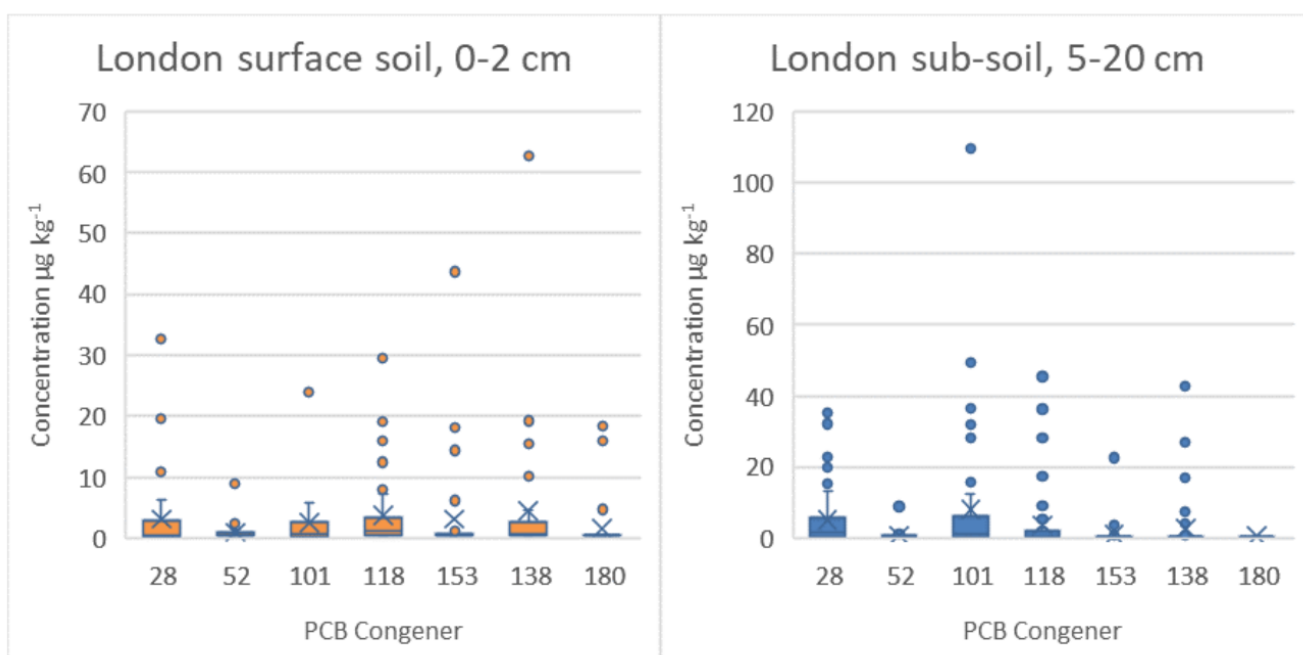
Figure 11 Box and whisper plot of PCB concentrations for London soils; displaying outliers (circle markers), highest and lowest non-outliers (upper and lower whisker limits), upper and lower quartiles (upper and lower box limits).

Individual PCB congeners together with associated homolog groups (tri to deca-chlorinated) can theoretically fingerprint different formulations (Arolclors by Monsanto, USA; Clophen by Bayer, Germany; Fenclor by Prodelec, France used in a wide variety of manufactured products (e.g. capacitors, transformers, building sealants, hydraulics, lubricants, resins, carbonless paper, paints, adhesives, inks, extenders and cutting oils). However, studies have shown that volatilization, partitioning, photo-degradation and microbial de-chlorination can confound direct source-environmental mixture matching [2, 7, 102, 103]. Given the differences in transport routes, weathering effects, absence of reference congener pattern and input mixtures as well as low numbers of congeners measured herein we consider an environmental forensic source apportionment beyond the scope of this baseline evaluation of organic pollution.

From a human health standpoint the general order of relative PCB toxic potency-biological activity proceeds according to substitution at the ortho position and stereochemical configuration such that: non-*ortho* planar PCBs > non-*ortho* planar PCBs > mono-*ortho* planar PCBs > di-*ortho* planar PCBs [104]. In this current study the most toxic PCB determined in the ICES-7 suite is PCB 118 which is one of eight mono-*ortho* (PCB 105, 114, 118, 123, 156, 157, 167, and 189) that can present an almost planar configuration similar to dioxins conferring a similar toxicity and biological effect. It is therefore important to note that PCB 118 is based on mean average the second most abundant of the ICES-7 congeners in the surface soils of London. The high abundance of PCB 118 in London soils relative to other congeners is not entirely unexpected given that it is an important constituent of unaltered Aroclor 1280, 1254, 1248, 1242 and 1016 [104]. Although the other ortho-PCBs (PCB 180, 153, 138, 118, 52, 28) presented here have a low affinity for the aryl hydrocarbon receptor, they are reported to cause a range of negative health effects including disruption to the endocrine system, neurotoxicity, immune suppression and inflammation [19].

3.7 Spatial Modelling of Organic Pollutants

Before proceeding with a discussion of spatial modelling of organic pollutants it is useful to further discuss the approach used herein. Inverse distance weighted (IDW) interpolation explicitly makes the assumption that samples that are close to one another are more alike than those that are further apart [105]. To predict a value for any unmeasured location, IDW uses the measured values surrounding the prediction location. The measured values closest to the prediction location have more influence on the predicted value than those farther away. IDW assumes that each measured point has a local influence that diminishes with distance. It gives greater weights to points closest to the prediction location, and the weights diminish as a function of distance, hence the name inverse distance weighted [105]. IDW predictions require 2 parameters the inverse distance power (p) and the number of nearest neighbours (n) to use. Whilst the premise behind IDW is broadly correct the choice of p and n is subjective and there is no relation between the prediction and the actual spatial variability as in other methods e.g. in kriging the model relates to the spatial variance of the parameter of interest through the variogram. In this instance we have taken a series of IDW predictors with varying combinations of p and n and use these as the

predictor covariates and the Random Forest model combines these covariates to model the estimated soil parameter at the prediction locations.

Model-predicted geochemical maps are presented for surface and subsurface soil TOC, BC, TPH and Σ^{16}PAH as well as benzo[a]pyrene in Figure 12 and Figure 13. These all to a lesser or greater extent confirm a gradual rise in the selected organic pollutants across the study area trending SW (low) to NE (High) or E (low) to W (high). Inspection of the predicted surface and sub surface soil TOC maps reveals only minor spatial differences (Figure 12). Similarly, the TPH maps show broadly comparable 2D spatial trends, as well as the effect of high measured values at Chelsea (site 16) (Figure 1 and Figure 11). The suggested W-E directionality of the soil pollution has been previously reported for the toxic metal mercury (total Hg) a surrogate for industrial activity/power generation/coal in sediments in the River Thames which reported low concentrations at sites in the west of London (e.g. Brentford, Chiswick, Barnes) and maximal concentrations in east London (Deptford, Millwall, O^2 arena, Barking, Woolwich) [1]. When taken together it could be speculated that the 2D soil pollution distribution might be transferred via urban tributaries to the River Thames.
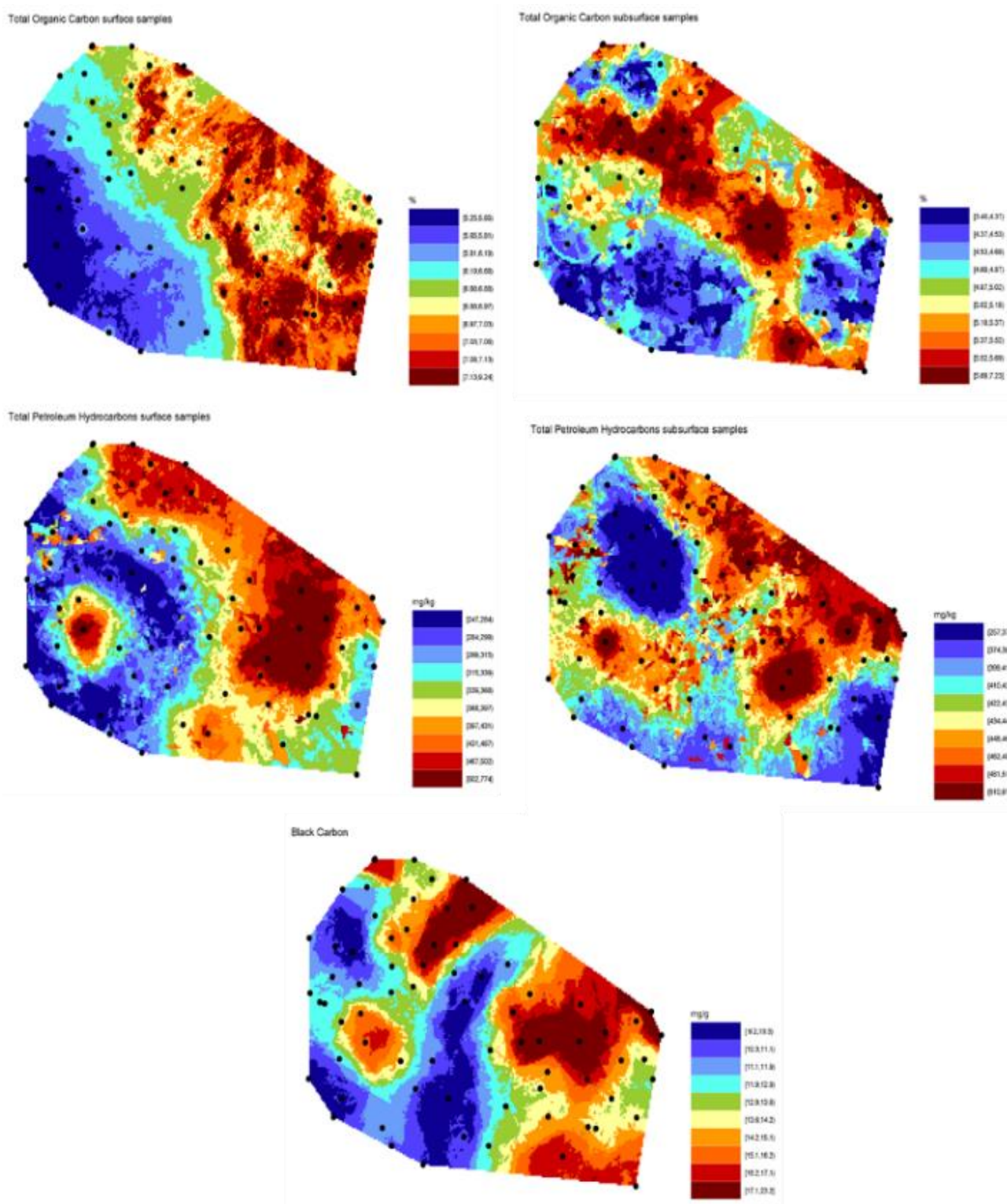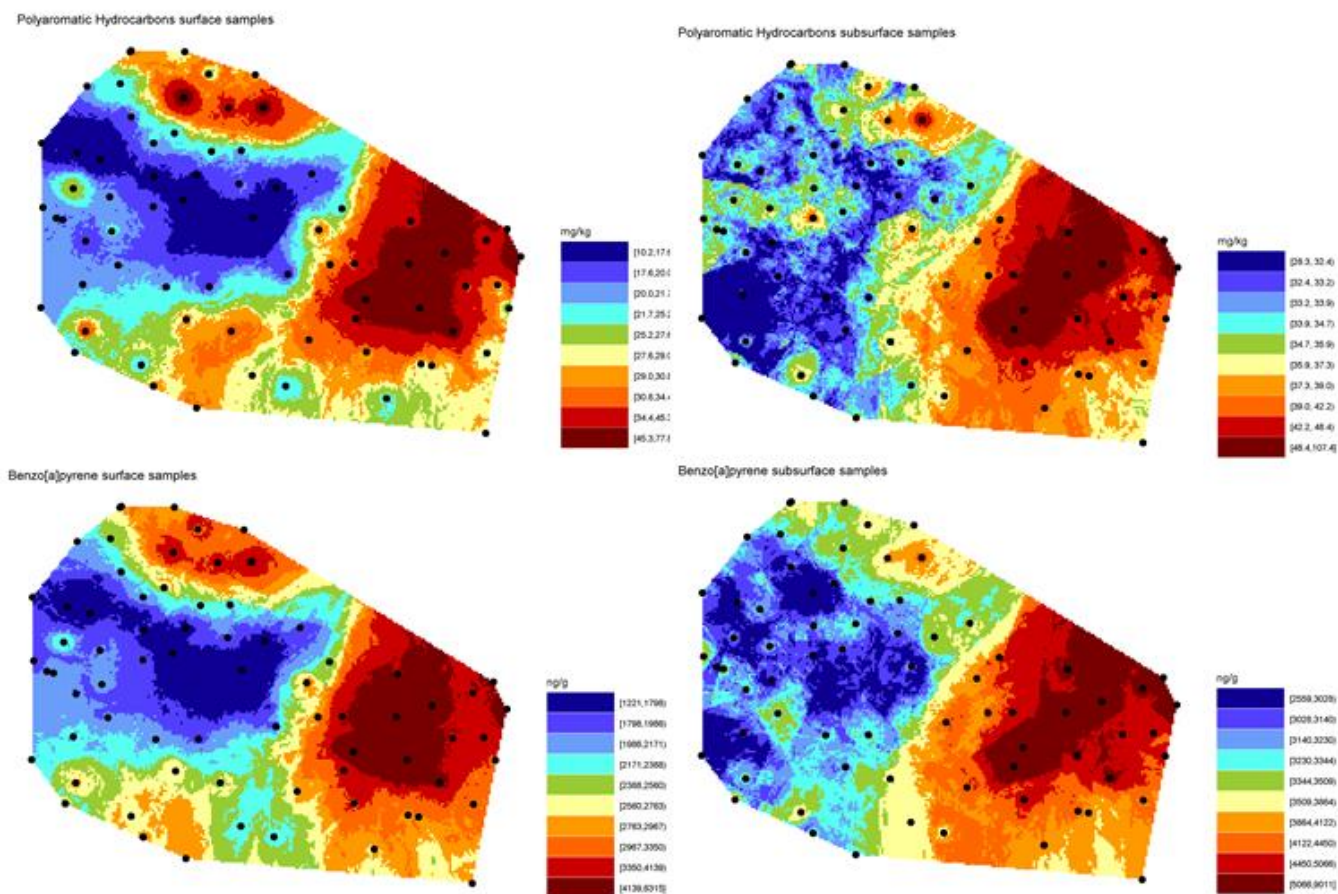
Figure 12 Interpolated map of TOC, TPH and BC in London soils.

Figure 13 Interpolated maps for summed PAH concentrations and for the benzo[*a*]pyrene in London soils.

In an urban environment such as London the spatial distribution of POPs is most likely to be driven by the anthropogenic activities according to land use [2, 28]. London has a complex history of land use [106] which is difficult to categorise. A number of studies, however, have shown that socio economic indicators can be a good proxy for land use [107-109] with the general premise that more industrialised land is associated with higher social deprivation. In this instance we have used the open source data on UK deprivation in 2019 provided by the Office of National Statistics [110] as additional predictor covariates for the RF modelling to provide the RF model an additional link to land use. Tables S2 to S5 show the significant predictor variables for BC, TOC, TPH, PAH and BaP respectively. The relative importance of each predictor (measured in the RF model by the gini-index [38]) are reported in decreasing order. The predictor variables beginning with "P" are IDW derived predictor variables with the number after the P being the inverse power and the number after the "N" is the number of nearest neighbours.

Four of the deprivation measures were identified as important for TOC (Table S3) and TPH (Table S4). For TOC only the subsurface samples included the deprivation measures of "Income" (measures the proportion of the population experiencing deprivation relating to low income [110]) and Wider ("wider barriers" which includes issues relating to access to housing such as affordability [110]). For TPH in the surface samples (Table S4) the deprivation indices "Income" , "Emplymnt" (the proportion of the working-age population in an area involuntarily excluded from the labour market [110]) and "Inc.opt" (measures the proportion of all those aged 60 or over who

experience income deprivation [110]). For TPH in the subsurface soils (Table S4) both the deprivation indices "Inc.opt" and "Income" are important predictors. It is important to note that we do not assert that TOC or TPH are in anyway a driver of income, employment or housing socio economic indicators only that there is a statistically significant link – which requires a further investigation. Furthermore, none of the deprivation indices were identified as being important predictor of total PAH or benzo[*a*]pyrene (Table S5-S6) which is probably best explained by the relatively recent timing of the socio economic study (2019) as compared to long history of PAH input [110].

3.8 Normal Background Concentrations and Practical Implications

Soil concentrations for selected PAH and PCB congeners were used to create a cautious upper estimate of concentrations of the population distribution taken across the study area (n=≤73, 5-20 cm interval); these compounds were selected as they are hazardous to human health. From a human health risk assessment perspective, the deeper samples also provide a more conservative estimate of the recorded concentrations (i.e. higher). The approach used was developed by the BGS for the UK Department of Environment Food and Rural Affairs (Defra) for deriving Normal Background Concentration (NBCs) to support decisions under the Environmental Protection Act 1990 in England and Wales but not the planning system [111]. The NBC method calculates the Upper Confidence Limit (UCL) of the 95th percentile (%ile) of the data. It assumes that data are of Gaussian distribution and if they are not then transformed data meet this criterion. The method assumes the data are representative of 'continuous fixed variation' from diffuse anthropogenic and geogenic inputs, excluding point sources. The UCL of the 95%ile limit of the data asserts that the majority of soil contaminant concentrations in an area will be below this value, point source pollution notwithstanding. To be clear the UCL of the 95%ile is an upper estimate of the upper end of the range of contaminant concentrations present either due to natural processes or diffuse pollution. It should not be confused with the 95% Upper Confidence Level (UCL95) of the population mean that is sometimes used as a representative contaminant concentration in an averaging area for comparison against a generic assessment criterion. The UCL95 is best described as a value that the true but unknown population mean contaminant concentration will be less than, with a 95% confidence.

The Defra NBC approach considers whether defining geographical and geochemical domains with distinctly different characteristics is appropriate for the sampled area. In this study we considered whether areas that might reasonably be expected to be ostensibly free of contamination e.g. parks (n=16) and domestic gardens (n=15) differ from the overall urban London area. When we examined the meta-data collected during sampling we found that the ground conditions described for parks and domestic gardens included descriptions of brick, slag, ceramic and glass (Supporting Information). These materials strongly indicate the anthropogenic component of soil composition and the potential for contamination. For this reason, we considered the entire study area dataset when calculating NBCs (n=73). Study area NBCs for selected PAH were calculated using the deeper 5-20 cm samples (Table 2). Study area NBCs for ICES ∑7 PCBs and PCB118 were also calculated using data for the deeper 5-20 cm samples. The NBC for the ICES ∑7PCBs is 110 µg.kg^{-1} and for PCB118 it is 59 µg.kg^{-1}. PCB118 was selected

because it is classed as a dioxin-like mono-ortho congener of potential risk to human health (Environment Agency, 2009).

Table 2 Study area NBCs for PAH compared with PAH generic assessment criteria.

| PAH | Study area NBC (mg.kg^{-1}) | Ratio of NBC to LQM CIEH S4ULs (6% SOM) | | | |
		Land use Residential with home grown produce	Allotment	Public open space near residential housing	Commercial & Industrial
Naphthalene	1.5	-	- -	- - - -	- - - -
Acenaphthylene	0.65	- - - -	- - -	- - - - -	- - - - - -
Acenaphthene	0.36	- - - -	- - -	- - - - -	- - - - -
Fluorene	1.5	- - - -	- - -	- - - - -	- - - - -
Phenanthrene	15	- -	- -	- - -	- - - -
Anthracene	3.9	- - - -	- - -	- - - - -	- - - - - -
Fluoranthene	38	- -	- -	- - -	- - - -
Pyrene	31	- - -	- -	- - -	- - - -
Benz[*a*]anthracene	16	+	+	-	- -
Chrysene	12	-	-	- -	- -
Benzo[*b*]fluoranthene	12	+	+	-	-
Benzo[*a*]pyrene	17	+	+	+	-
Indeno[*1,2,3-cd*]pyrene	11	-	-	- -	- - -
Dibenz[*a,h*]anthracene	1.7	+	+	+	-
Benzo[*ghi*]perylene	10	- -	- - -	- - -	- - -
∑USEPA 16 PAH	80	Na	na	na	na

NOTES: na = not applicable; S4ULs [112]. Land Quality Management © 2015

NBC is less than the S4UL by less than: 1 order of magnitude (OOM) (-); 2 OOM (--); 3 OOM (---);

NBC is more than S4UL by less than 1 OOM (+).

In order to evaluate the possible implications for risk-based land management decisions under the contaminated land regime in Part 2A of the Environmental Protection Act 1990 (Part 2A), the NBCs for PAH congeners were compared to published Generic Assessment Criteria (GAC) developed by LQM [112]. The NBC method is specifically not to be used for land use planning purposes. Concentrations below the NBC would in general not indicate an elevated level of risk to human health and would not result in regulatory action under Part 2A. The LQM CIEH Suitable for Use Levels (S4UL) are soil contaminant concentrations that for a specific land use scenario would give rise to an estimated Average Daily Exposure equal to Health Criteria Values representing a minimal (non threshold behaviour) or no appreciable (threshold behaviour) health risk. Such land use scenarios could be a female child living in a home with a garden that may be used for home grown produce or an adult female working in a commercial building.

England's National Planning Policy Framework states that land subject to a change in planning permission must be safe and "suitable for its proposed use taking account of ground conditions

and any risks arising from land instability and contamination" [113]. An assessment of suitability for use is informed by comparing site specific contaminant concentrations representative of the area being investigated with relevant GAC. Table 2 compares study area NBCs for selected PAH with the LQM CIEH Suitable for Use Levels (S4UL) GACs for four land uses (residential with plant uptake, allotments, public open space and commercial & industrial). The GACs selected were 6% Soil Organic Matter (SOM); TOC was recorded between 1.75 and 9.61 with a mean of 5.06 %, using a conversion factor of 1.74 showed SOM proportions range between 3.05% and 16.7% [114].

Most PAH do not exceed the S4ULs for any of the land uses. The comparison however shows that the study area NBCs for BaA, BaP, BbF and DBA exceed S4ULs for three land uses: residential with homegrown produce, allotments and public open space near residential housing. This could trigger further inspection of land under Part 2A. It could also be seen as an indication that should a change of land use be required for any reason then further investigation of the potential soil contamination are likely to be required on a site-specific basis to ensure the land is safe and suitable for the intended use. Any assessment would need to be consistent with regulatory requirements such as those explained in the Environment Agency Land Contamination Risk Management guidance [115].

The NBC for PCB 118 (59 µg.kg^{-1}) was compared with Soil Guideline Values (SGV) for dioxins, furans and dioxin-like PCBs (residential: 8 µg.kg^{-1}; allotments: 8 µg.kg^{-1}; and commercial: 240 µg.kg^{-1}) [78]. SGVs are GAC published by the Environment Agency. Most have been superceded by other GACs but those for dioxins, furans and dioxin-like PCBs remain current and relevant. The congener distribution for this study is assumed to be sufficiently similar to that used in the derivation of the SGVs, which itself is based on diffuse signatures measured by the UK soil and herbage pollutant survey (Environment Agency 2007). The PCB118 NBC exceeds the SGVs for residential and allotment scenarios. The exceedances are subject to the same assumptions made for PAH where further inspection of the land may be triggered under different regulatory regimes in England in order to assess risk to human health. It should be noted that exceedance of the SGVs does not account for concentrations of dioxins, furans and other dioxin-like PCB congeners that fall within the scope of the SGVs as these have not been measured. The overall approach taken here to evaluating the PCB data in the context of human health is consistent with earlier work by the authors [2]. The early study reported a lower PCB118 NBC (4.4 µg.kg^{-1}) for a different and smaller area of south east London.

4. Conclusions

1) To the best of our knowledge this is one of the largest surveys of organic pollutants in London soils. This baseline study highlights both the variation in concentrations and commonality of source inputs across a range of urban domains. Both PAH and PCB pollutants were observed at every location and at both depth intervals confirming their ubiquity, persistence and potential to interact with humans and ecology.

2) London surface soils contain higher amounts of anthropogenic BC than most cities, suggesting continued atmospheric deposition from traffic and possibly incorporation of older burnt wastes from legacy activities and industries. The presence of BC at high concentrations is not entirely unexpected given that BC is largely thought to cycle (decompose) extremely slowly. However, it should be borne in mind that the effect of soil BC on soil ecosystem health is probably

driven as much by the form of BC molecular structures present (and their ability to sorb metal and organic contaminants) as the actual concentration.

3) Soils were evaluated for TPH and PAH in order to provide a baseline hydrocarbon chemistry for site owners, occupiers and regulators involved in understanding the potential risk to human health and the environment. The ubiquity of target organic pollutants in London soils was confirmed by the presence of petroleum hydrocarbons at all sites at both depths; of the 146 soils analysed 33 exceeded a non-regulatory benchmark of ~500 mg.kg^{-1}. TPH pollution was significantly lower at surface than in the sub soils suggesting a general decrease in TPH pollution over time. London soil PAH concentrations (Σ^{16}PAH 29.51, median 19.00 mg.g^{-1}; Σ^{27}PAH mean 39.29, median 25.64 mg.g^{-1}) were higher than many other global cities suggesting either greater PAH input or possibly enhanced. Parent PAH ratios confirmed that both depth intervals were mainly sourced from combustion processes such as traffic input (vehicle exhaust, road-surface ware) and coal-wood-vegetation burning. However, the pattern of alkylated PAH (e.g. phenanthrenes) indicated minor petroleum/oil PAH input. We speculate that the lower PAH concentrations observed in the surface soils as compared to subsoils may be due a combination of improved environmental regulations (restrictions on burning of coal/ open domestic fires/ industrial furnaces (e.g. Clean Air Act 1956, 1968)), decommissioning of power stations across central London (1960s-1980s) as well as more recent vehicle congestion charging (2003- present) and low emissions zones.

4) Following a worldwide manufacturing ban in 1981, central London's soils contain a legacy background of PCB pollution at a mean of Σ^7 PCB 20.82 and median of 9.68 µg.kg^{-1} which is 10 to 16 times higher than equivalent rural soils. London sub soils contain higher PCB concentrations than surface soils which is commensurate with cessation of production and gradual disposal of manufactured goods containing PCB. However, their presence in surface soils clearly demonstrates their environmental persistence and continued release into the environment, presumably from secondary sources such as demolition dusts, leaking capacitors/transformer fluids.

5) Model-predicted geochemical maps presented for surface and subsurface soil TOC, BC, TPH and Σ^{16}PAH as well as benzo[*a*]pyrene indicate a general increase in concentrations along a W-E trend. These spatial patterns / trends are complex but broadly commensurate with the higher density of historical industry situated in east London. The link between present day social deprivation indices and POPs (TOC, BC, TPH, PAH) was explored and indicated that TOC and TPH were related to low Income, exclusion from the labour market and housing issues such as affordability. However, the precise reason for this association is not easy to interpret meaningfully.

6) NBCs calculated for BaA, BaP and DBA exceeded the risk-based S4UL human health generic assessment criteria for three land uses (residential with homegrown produce, allotments and public open space near residential housing). The comparison shows that for these compounds this could trigger further inspection of land under Part 2A. It could also be seen as an indication that should a change of land use be required for any reason then further investigation of the potential soil contamination are likely to be required on a site-specific basis to ensure the land is safe and suitable for the intended use. Similarly, NBCs calculated for dioxin-like PCB118 exceeded the soil guide line value for residential and allotments which could in turn trigger further inspection.

Acknowledgments

Additional Materials (if any)

The following additional materials are uploaded at the page of this paper.

1. Methods - Total Petroleum Hydrocarbons.
2. Methods - Spatial Modelling and Interpolation – stages i-iv
3. Figure S1: London soil sampling points with a convex hull drawn round the points to define a prediction region and a prediction grid set out on 500 m squares.
4. Table S1: t-test results comparing surface (0-2 cm) and subsurface (5-20 cm) geochemistry.
5. Table S2: Significant predictor values for black carbon (BC).
6. Table S3: Significant predictor values for total organic carbon (TOC).
7. Table S4: Significant predictor values for total petroleum hydrocarbons (TPH).
8. Table S5: Significant predictor values for benzo[a]pyrene.
9. Table S6: Significant predictor values for polycyclic aromatic hydrocarbons (PAH).

Author Contributions

C.H. Vane conceived the study, managed the analyses and interpreted the data, A.W. Kim acquired the PAH data, S. Lowe acquired the PCB data, R.A.L dos Santos assisted in acquiring the BC data, V. M-Hayes acquired TOC and TPH and BC data, D.J. Beriro, M.R. Cave analysed the data (statistically) and C. Collins revised the manuscript. Paul Nathanail contributed the comparison with the LQM CIEH S4ULs and ensured the text reflected relevant legal contexts and technical guidance.

Funding

Competing Interests

The authors have declared that no competing interests exist.

References

1. Vane CH, Beriro DJ, Turner GH. Rise and fall of mercury (Hg) pollution in sediment cores of the Thames Estuary, London, UK. Earth Environ Sci Trans R Soc Edinb. 2015; 105: 285-296.
2. Vane CH, Kim AW, Beriro DJ, Cave MR, Knights K, Moss-Hayes V, et al. Polycyclic aromatic hydrocarbons (PAH) and polychlorinated biphenyls (PCB) in urban soils of Greater London, UK.

Appl Geochemistry. 2014; 51: 303-314.

3. Vane CH, Turner GH, Chenery SR, Richardson M, Cave MR, Terrington R, et al. Trends in heavy metals, polychlorinated biphenyls and toxicity from sediment cores of the inner River Thames estuary, London, UK. Environ Sci Processes Impacts. 2020; 22: 364-380.

4. Appleton JD, Cave MR. Variation in soil chemistry related to different classes and eras of urbanisation in the London area. Appl Geochemistry. 2018; 90: 13-24.

5. London Environment Strategy [Internet]. London: Greater London Authority; 2018. Available from: https://www.london.gov.uk/sites/default/files/london_environment_strategy_0.pdf.

6. Cave MR, Wragg J, Beriro DJ, Vane C, Thomas R, Riding M, et al. An overview of research and development themes in the measurement and occurrences of polyaromatic hydrocarbons in dusts and particulates. J Hazard Mater. 2018; 360: 373-390.

7. Vane CH, dos Santos RA, Kim AW, Moss-Hayes V, Fordyce FM, Bearcock JM. Peristent organic pollutants (PAH, PCB, TPH) in freshwater, urban tributary and estuarine surface sediments of the River Clyde, Scotland, UK. Earth Environ Sci Trans R Soc Edinb. 2019; 108: 299-314.

8. Vane CH, Kim AW, Moss-Hayes V, Turner G, Mills K, Chenery SR, et al. Organic pollutants, heavy metals and toxicity in oil spill impacted salt marsh sediment cores, Staten Island, New York City, USA. Mar Pollut Bull. 2020; 151: 110721.

9. Peters KE, Walters CC, Moldowan JM. The biomarker guide. Cambridge: Cambridge University Press; 2005.

10. Stout SA, Wang Z. Oil spill environmental forensics: Fingerprinting and source identification. Amsterdam: Elsevier; 2007.

11. Thomas MJ, Collinge E, Witt M, Palacio Lozano DC, Vane CH, Moss-Hayes V, et al. Petroleomic depth profiling of Staten Island salt marsh soil: 2ω detection FTICR MS offers a new solution for the analysis of environmental contaminants. Sci Total Environ. 2019; 662: 852-862.

12. Wilcke W. Polycyclic aromatic hydrocarbons (PAHs) in soil - a review. J Plant Nutr Soil Sci. 2000; 163: 229-248.

13. Gale SL, Noth EM, Mann J, Balmes J, Hammond SK, Tager IB. Polycyclic aromatic hydrocarbon exposure and wheeze in a cohort of children with asthma in Fresno, CA. J Expo Sci Environ Epidemiol. 2012; 22: 386-392.

14. Deziel NC, Rull RP, Colt JS, Reynolds P, Whitehead TP, Gunier RB, et al. Polycyclic aromatic hydrocarbons in residential dust and risk of childhood acute lymphoblastic leukemia. Environ Res. 2014; 133: 388-395.

15. Dore CJ, Watterson JD, Murrells TP, Passant NR, Hobson MM, Baggott SL, et al. UK emissions of air pollutants 1970 to 2003. Culham, Oxfordshire: Atomic Energy Authority; 2005; AEAT/ENV/R-2160.

16. Harrad S, Stuart AP, Alcock R, Boumphrey R, Burnett V, Duarte-Davidson R, et al. Polychlorinated biphenyls (PCBs) in the British environment: Sinks, sources and temporal trends. Environ Pollut. 1994; 85: 131-146.

17. Beyer A, Biziuk M. Environmental fate and global distribution of polychlorinated biphenyls. In: Reviews of environmental contamination and toxicology (continuation of residue reviews). Boston, MA, USA: Springer; 2009.

18. The stockholm convention on persistent organic pollutants [Internet]. Stockholm, Sweden: Stockholm Convention; 2004. Available from: http://chm.pops.int/portals/0/repository/convention_text/unep-pops-cop-convtext-

full.english.pdf.

19. Abella V, Santoro A, Scotece M, Conde J, López-López V, Lazzaro V, et al. Non-dioxin-like polychlorinated biphenyls (PCB 101, PCB 153 and PCB 180) induce chondrocyte cell death through multiple pathways. Toxicol Lett. 2015; 234: 13-19.

20. Colter BT, Garber HF, Fleming SM, Fowler JP, Harding GD, Hooven MK, et al. Ahr and Cyp1a2 genotypes both affect susceptibility to motor deficits following gestational and lactational exposure to polychlorinated biphenyls. Neurotoxicology. 2018; 65: 125-134.

21. Minasny B, McBratney AB. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. Geoderma. 2007; 140: 324-336.

22. Zhu Q, Lin HS. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. Pedosphere. 2010; 20: 594-606.

23. Hengl T, Nussbaum M, Wright MN, Heuvelink GB, Gräler B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. Peer J. 2018; 6: e5518.

24. Kirkwood C, Cave M, Beamish D, Grebby S, Ferreira A. A machine learning approach to geochemical mapping. J Geochem Explor. 2016; 167: 49-61.

25. Sekulić A, Kilibarda M, Heuvelink G, Nikolić M, Bajat B. Random forest spatial interpolation. Remote Sens. 2020; 12: 1687.

26. Cave MR. A machine learning approach to geostatistics applied to contaminants in soil 6D.2. Proceedings of the CleanUp 2017 Conference; 2017 September 10-14; Melbourne, Australia.

27. Kim A, Vane C, Moss-Hayes V, Beriro D, Nathanail C, Fordyce F, et al. Polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) in urban soils of Glasgow, UK. Earth Environ Sci Trans R Soc Edinb. 2019; 108: 231-247.

28. Vane CH, Harrison I, Kim AW. Polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) in sediments from the Mersey Estuary, U.K. Sci Total Environ. 2007; 374: 112-126.

29. Vane CH, Kim AW, Emmings JF, Turner GH, Moss-Hayes V, Lort JA, et al. Grain size and organic carbon controls polyaromatic hydrocarbons (PAH), mercury (Hg) and toxicity of surface sediments in the River Conwy Estuary, Wales, UK. Mar Pollut Bull. 2020; 158: 111412.

30. Beriro DJ, Vane CH, Cave MR, Nathanail CP. Effects of drying and comminution type on the quantification of Polycyclic Aromatic Hydrocarbons (PAH) in a homogenised gasworks soil and the implications for human health risk assessment. Chemosphere. 2014; 111: 396-404.

31. Vane CH, Rawlins BG, Kim AW, Moss-Hayes V, Kendrick CP, Leng MJ. Sedimentary transport and fate of polycyclic aromatic hydrocarbons (PAH) from managed burning of moorland vegetation on a blanket peat, South Yorkshire, UK. Sci Total Environ. 2013; 449: 81-94.

32. Rawlins BG, Vane CH, Kim AW, Tye AM, Kemp SJ, Bellamy PH. Methods for estimating types of soil organic carbon and their application to surveys of UK urban areas. Soil Use Manag. 2008; 24: 47-59.

33. Kuhlbusch TA. Method for determining black carbon in residues of vegetation fire. Environ Sci Technol. 1995; 29: 2695-2702.

34. Vane CH, Harrison I, Kim AW. Assessment of polyaromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) in surface sediments of the Inner Clyde Estuary, UK. Mar Pollut Bull. 2007; 54: 1301-1306.

35. Vane CH, Harrison I, Kim AW, Moss-Hayes V, Vickers BP, Hong K. Organic and metal

contamination in surface mangrove sediments of south China. Mar Pollut Bull. 2009; 58: 134-144.

36. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: Foundation for Statistical Computing; 2020.

37. Pebesma E. Simple features for R: Standardized support for spatial vector data. R J. 2018; 10: 439-446.

38. Wright MN, Ziegler A. Ranger: A fast implementation of random forests for high dimensional data in C++ and R. J Stat Softw. 2017; 77: 1-17.

39. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw. 2010; 36: 1-13.

40. Cooper HV, Vane CH, Evers S, Aplin P, Girkin NT, Sjögersten S. From peat swamp forest to oil palm plantations: The stability of tropical peatland carbon. Geoderma. 2019; 342: 109-117.

41. Girkin NT, Vane CH, Cooper H, Moss-Hayes V, Craigon J, Turner BL, et al. Spatial variability of organic matter properties determines methane fluxes in a tropical forested peatland. Biogeochemistry. 2019; 142: 231-245.

42. Devoy RJ. Analysis of the geological evidence for Holocene sea-level movements inb southeast England. Proc Geol Assoc. 1982; 93: 65-90.

43. Khan SN, Vane CH, Horton BP, Hillier C, Riding JB, Kendrick C. The application of δ13C, TOC and C/N geochemistry to reconstruct Holocene relative sea levels and paleoenvironments in the Thames Estuary, UK. J Quat Sci. 2015; 30: 417-433.

44. Schumacher BA. Methods for the determination of total organic carbon (TOC) in soils and sediments [Internet]. USA: US Environment Protection Agency, Office of Research and Development; 2002. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.587.494&rep=rep1&type=pdf.

45. Hoogsteen MJ, Lantinga EA, Bakker EJ, Groot JC, Tittonell PA. Estimating soil organic carbon through loss on ignition: Effects of ignition conditions and structural water loss. Eur J Soil Sci. 2015; 66: 320-328.

46. Bradley RI, Milne R, Bell J, Lilly A, Jordan C, Higgins A. A soil carbon and land use database for the United Kingdom. Soil Use Manag. 2005; 21: 363-369.

47. Verheijen FG, Bellamy PH, Kibblewhite MG, Gaunt JL. Organic carbon ranges in arable soils of England and Wales. Soil Use Manag. 2005; 21: 2-9.

48. Reynolds B, Chamberlain PM, Poskitt J, Woods C, Scott WA, Rowe EC, et al. Countryside survey: National "soil change" 1978–2007 for topsoils in Great Britain—acidity, carbon, and total nitrogen status. Vadose Zone J. 2013; 12: 1-15.

49. Pouyat R, Groffman P, Yesilonis I, Hernandez L. Soil carbon pools and fluxes in urban ecosystems. Environ Pollut. 2002; 116: S107-S118.

50. Beyer L, Kahle P, Kretschmer H, Wu QG. Soil organic matter composition of man-impacted urban sites in North Germany. J Plant Nutr Soil Sci. 2001; 164: 359-364.

51. Beesley L. Carbon storage and fluxes in existing and newly created urban soils. J Environ Manag. 2012; 104: 158-165.

52. Vasenev V, Kuzyakov Y. Urban soils as hot spots of anthropogenic carbon accumulation: Review of stocks, mechanisms and driving factors. Land Degrad Dev. 2018; 29: 1607-1622.

53. Raciti SM, Hutyra LR, Finzi AC. Depleted soil carbon and nitrogen pools beneath impervious surfaces. Environ Pollut. 2012; 164: 248-251.

54. Jo HK. Impacts of urban greenspace on offsetting carbon emissions for middle Korea. J Environ Manag. 2002; 64: 115-126.

55. Baldock JA, Masiello CA, Gelinas Y, Hedges JI. Cycling and composition of organic matter in terrestrial and marine ecosystems. Mar Chem. 2004; 92: 39-64.

56. Edmondson JL, Stott I, Potter J, Lopez-Capel E, Manning DA, Gaston. KJ, et al. Black carbon contribution to organic carbon stocks in urban soil. Environ Sci Technol. 2015; 49: 8339-8346.

57. Liu SD, Xia XH, Zhai YW, Wang R, Liu T, Zhang SW. Black carbon (BC) in urban and surrounding rural soils of Beijing, China: Spatial distribution and relationship with polycyclic aromatic hydrocarbons (PAHs). Chemosphere. 2011; 82: 223-228.

58. Agarwal T, Bucheli TD. Is black carbon a better predictor of polycyclic aromatic hydrocarbon distribution in soils than total organic carbon? Environ Pollut. 2011; 159: 64-70.

59. Azzolina NA, Kreitinger JP, Skorobogatov Y, Shaw RK. Background concentrations of PAHs and metals in surface and subsurface soils collected throughout Manhattan, New York. Environ Forensics. 2016; 17: 294-310.

60. He Y, Zhang GL. Historical record of black carbon in urban soils and its environmental implications. Environ Pollut. 2009; 157: 2684-2688.

61. Bucheli TD, Blum F, Desaules A, Gustaffson Ö. Polycyclic aromatic hydrocarbons, black carbon, and molecular markers in soils of Switzerland. Chemosphere. 2004; 56: 1061-1076.

62. Nam JJ, Gustafsson O, Kurt-Karakus P, Breivik K, Steinnes E, Jones KC. Relationships between organic matter, black carbon and persistent organic pollutants in European background soils: Implications for sources and environmental fate. Environ Pollut. 2008; 156: 809-817.

63. Power stations on the River Thames, shipyards on the River Thames. In: Industry on the River Thames. 1st ed. Memphis, Tennessee, USA: Books LLC; 2010.

64. Lohmann R, MacFarlane JK, Gschwend PM. Importance of black carbon to sorption of native PAHs, PCBs, and PCDDs in Boston and New York, Harbor sediments. Environ Sci Technol. 2005; 39: 141-148.

65. Koelmans AA, Jonker MT, Cornelissen G, Bucheli TD, Van Noort PC, Gustafsson O. Black carbon: The reverse of its dark side. Chemosphere. 2006; 63: 365-377.

66. Yuan CP, Gao BL, Peng YT, Gao X, Fan BB, Chen Q. A meta-analysis of heavy metal bioavailability response to biochar aging: Importance of soil and biochar properties. Sci Total Environ. 2021; 756: 144058.

67. Steiner C, Teixeira WG, Lehmann J, Nehls T, de Macêdo JL, Blum WE, et al. Long term effects of manure, charcoal and mineral fertilization on crop production and fertility on a highly weathered Central Amazonian upland soil. Plant Soil. 2007; 291: 275-290.

68. Nwokoro C, Ewin C, Harrison C, Ibrahim M, Dundas I, Dickson I, et al. Cycling to work in London and inhaled does of black carbon. Eur Respir J. 2012; 40: 1091-1097.

69. Vane CH, Kim AW, Cave MR, Li N. Equal abundance of odd and even n-alkanes from cycad leaves: Can the carbon preference index (CPI) faithfully record terrestrial organic matter input at low lattitudes. Encephalartos. 2014; 118: 10-12.

70. Zhang ZH, Zhao MX, Yang XD, Wang SM, Jiang XZ, Oldfield F, et al. A hydrocarbon biomarker record for the last 40 kyr of plant input to Lake Heqing, southwestern China. Org Geochem. 2004; 35: 595-613.

71. Mills K, Vane CH, dos Santos RA, Ssemmanda I, Leng MJ, Ryves D. Linking land and lake: Using novel geochemical techniques to understand biological response to environmental change.

Quat Sci Rev. 2018; 202: 122-138.

72. Vane CH, Harrison I, Kim AW, Moss-Hayes V, Vickers BP, Horton BP. Status of organic pollutants in surface sediments of Barnegat Bay-Little Egg Harbor Estuary, New Jersey, USA. Mar Pollut Bull. 2008; 56: 1802-1814.

73. Bradley LJ, Magee BH, Allen SL. Background levels of polyaromatic hydrocarbons (PAH) and selected metals in New England urban soils. J Soil Contam. 1994; 3: 349-361.

74. Volkman JK, Holdsworth DG, Neill GP, Bavor HJ. Identification of natural, anthropogenic and petroleum hydrocarbons in aquatic sediments. Science Total Environ. 1992; 112: 203-219.

75. Simpson SL, Batley GE, Charlton AA. Revision of the ANZECC/ARMCANZ sediment quality gudielines. Lucas Heights, NSW Australia: CSIRO; 2013.

76. Napolitano GE, Richmond JE, Stewart AJ. Characterization of petroleum-contaminated soils by thin-layer chromatography with flame ioniozation detection. J Soil Contam. 1998; 7: 709-724.

77. Vane CH, Chenery SR, Harrison I, Kim AW, Moss-Hayes VM, Jones DG. Chemical signatures of the Anthropocene in the Clyde estuary, UK: Sediment-hosted Pb, 207/206Pb, total petroleum hydrocarbons, polyaromatic hydrocarbon and polychlorinated biphenyl pollution records. Philos Trans Royal Soc A. 2011; 369: 1085-1111.

78. Creaser CS, Wood MD, Alcock RE, Copplestone D, Crook PJ, Barraclogh D. UKSHS report No. 9 environmental concentrations of polycyclic aromatic hydrocarbons in UK soil and herbage. Bristol, England: Environment Agency, 2007.

79. Agarwal T. Concentration level, pattern and toxic potential of PAHs in traffic soil of Delhi, India. J Hazard Mater. 2009; 171: 894-900.

80. Peng C, Chen WP, Liao XL, Wang ME, Ouyang ZY, Jiao WT, et al. Polycyclic aromatic hydrocarbons in urban soils of Beijing: Status, sources, distribution and potential risk. Environ Pollut. 2011; 159: 802-808.

81. Hindersmann B, Förster A, Achten C. Novel and specific source identification of PAH in urban soils: Alk-PAH-BPCA index and "V"-shape distribution pattern. Environ Pollut. 2020; 257: 113594.

82. Hedges JI, Keil RG. Organic geochemical perspectives on estuarine processes: Sorption reactions and consequences. Mar Chem. 1999; 65: 55-65.

83. Huang WL, Peng PA, Yu ZQ, Fu JM. Effects of organic matter heterogeneity on sorption and desorption of organic contaminants by soils and sediments. Appl Geochemistry. 2003; 18: 955-972.

84. Ukalska-Jaruga A, Smreczak B, Klimkowicz-Pawlas A. Soil organic matter composition as a factor affecting the accumulation of polycyclic aromatic hydrocarbons. J Soils Sediments. 2018; 19: 1890-1900.

85. MacDonald DD, Dipinto LM, Field J, Ingersoll CG, Long ER, Swartz RC. Development and evaluation of consensus-based sediment effect concentrations for polychlorinated biphenyls. Environ Toxicol Chem. 2000; 19: 1403-1413.

86. Tobiszewski M, Namiesnik J. PAH diagnostic ratios for the identification of pollution emission sources. Environ Pollut. 2012; 162: 110-119.

87. Yunker MB, Macdonald RW, Vingarzan RV, Mitchell RH, Goyette D, Sylvestre S. PAHs in the fraser river basin: A critical appraisal of PAH ratios as indicators of PAH source and composition. Org Geochem. 2002; 33: 489-515.

88. Costa HJ, Sauer TC. Forensic approaches and considerations in identifying PAH background.

Environ Forensics. 2005; 6: 9-16.

89. Boehm PD, Pietari J, Cook LL, Saba T. Improving rigor in polycyclic aromatic hydrocarbon source fingerprinting. Environ Forensics. 2018; 19: 172-184.

90. Yan BZ, Abrajano TA, Bopp RF, Chaky DA, Benedict LA, Chillrud SN. Molecular tracers of saturated and polycyclic aromatic hydrocarbon inputs into central park lake, New York City. Environ Sci Technol. 2005; 39: 7012-7019.

91. Saber D, Mauro D, Sirivedhin T. Environmental forensics investigation in sediments near a former manufactured gas plant site. Environ Forensics. 2006; 7: 65-75.

92. Neff JM, Stout SA, Gunster DG. Ecological risk assessment of polycyclic aromatic hydrocarbons in sediments: Identifying sources and ecological hazard. Integr Environ Assess Manag. 2009; 1: 22-33.

93. Sundahl M, Sikander E, Ek-Olausson B, Hjorthage A, Rosell L, Tornevall M. Determinations of PCB within a project to develop cleanup methods for PCB-containing elastic sealant used in outdoor joints between concrete blocks in buildings. J Environ Monit. 1999; 1: 383-387.

94. Zhou JL, Siddiqui E, Ngo HH, Guo WS. Estimation of uncertainty in the sampling and analysis of polychlorinated biphenyls and polycyclic aromatic hydrocarbons from contaminated soil in Brighton, UK. Sci Total Environ. 2014; 497-498: 163-171.

95. Leah R, Johnson M, Connor L, Fox W, Levene C. Dispersal of polychlorinated biphenyls from a closed landfill site. Land Contam Reclam. 2001; 9: 1-8.

96. Cachada A, Lopes LV, Hursthouse AS, Biasioli M, Grčman H, Otabbong E, et al. The variability of polychlorinated biphenyls levels in urban soils from five European cities. Environ Pollut. 2009; 157: 511-518.

97. Motelay-Massei A, Ollivon D, Garban B, Teil MJ, Blanchard M, Chevreuil M. Distribution and spatial trends of PAHs and PCBs in soils in the Seine River basin, France. Chemosphere. 2004; 55: 555-565.

98. Wilcke W, Krauss M, Safronov G, Fokin AD, Kaupenjohann M. Polychlorinated biphenyls (PCBs) in soils of the Moscow region: Concentrations and small-scale distribution along an urban–rural transect. Environ Pollut. 2006; 141: 327-335.

99. Yu HY, Liu YF, Shu XQ, Ma LM, Pan YW. Assessment of the spatial distribution of organochlorine pesticides (OCPs) and polychlorinated biphenyls (PCBs) in urban soil of China. Chemosphere. 2020; 243: 125392.

100. Kumar B, Verma VK, Singh SK, Kumar S, Sharma CS, Akolkar AB. Polychlorinated biphenyls in residential soils and their health risk and hazard in an industrial city in India. J Public Health Res. 2014; 3: 252.

101. Kim L, Jeon JW, Son JY, Park MK, Kim CS, Jeon HJ, et al. Monitoring and risk assessment of polychlorinated biphenyls (PCBs) in agricultural soil from two industrialized areas. Environ Geochem Health. 2017; 39: 279-291.

102. Sokol RC, Kwon OS, Bethoney CM, Rhee GY. Reductive dechlorination of polychlorinated biphenyls in st. Lawrence River sediments and variations in dechlorination characteristics. Environ Sci Technol. 1994; 28: 2054-2064.

103. Quensen JF, Tiedje JM, Boyd SA. Reductive dechlorination of polychlorinated biphenyls by anaerobic microorganisms from sediments. Science. 1988; 242: 752-754.

104. Rushneck DR, Beliveau A, Fowler B, Hamilton C, Hoover D, Kaye K, et al. Concentrations of dioxin-like PCB congeners in unweathered Aroclors by HRGC/HRMS using EPA Method 1668A.

Chemosphere. 2004; 54: 79-87.

105. Fortin MJ, Dale MR. Spatial analysis: A guide for ecologists. Cambridge: Cambridge University Press; 2005.

106. Sheppard FH. London: A history. Oxford, England: Oxford University Press; 1998.

107. Pauleit S, Ennos R, Golding Y. Modeling the environmental impacts of urban land use and land cover change—a study in Merseyside, UK. Landsc Urban Plan. 2005; 71: 295-310.

108. Pearce JR, Richardson EA, Mitchell RJ, Shortt NK. Environmental justice and health: The implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom. Trans Inst Br Geogr. 2010; 35: 522-539.

109. Su SL, Zhang QW, Pi JH, Wan C, Weng M. Public health in linkage to land use: Theoretical framework, empirical evidence, and critical implications for reconnecting health promotion to land use policy. Land Use Policy. 2016; 57: 605-618.

110. McLennan D, Noble S, Noble M, Plunkett E, Wright G, Gutacker N. The English Indices of Deprivation 2019: Technical report. London, England: Ministry of Housing, Communities and Local Government; 2019.

111. Ander EL, Johnson CC, Cave MR, Palumbo-Roe B, Nathanail CP, Lark RM. Methodology for the determination of normal background concentrations of contaminants in English soil. Sci Total Environ. 2013; 454-455: 604-618.

112. Nathanail CP, McCaffrey C, Gillett AG, Ogden RC, Nathanail JF. The LQM/CIEH S4ULs for human health risk assessment. Nottingham: LQM Press; 2015.

113. Ministry of Housing, Communities & Local Government. National Planning Policy Framework. London: Ministry of Housing, Communities & Local Government; 2019. p. 1-76. Available from: https://www.gov.uk/government/publications/national-planning-policy-framework--2.

114. Pribyl DW. A critical review of the conventional SOC to SOM conversion factor. Geoderma. 2010; 156: 75-83.

115. Land contamination risk management (LCRM) [Internet]. Bristol, England: Environment Agency; 2020 [cited 2020 October 8]. Available from: https://www.gov.uk/government/publications/land-contamination-risk-management-lcrm.

Enjoy *AEER* by:

1. Submitting a manuscript
2. Joining in volunteer reviewer bank
3. Joining Editorial Board
4. Guest editing a special issue

For more details, please visit:
http://www.lidsen.com/journals/aeer