Spatial Distribution Patterns of Chosen Chemical Elements with Different Spatial Scales in London Topsoils and Lake Taihu Sediments

Yuting Meng

Supervisor: Dr. Chaosheng Zhang

The thesis is submitted to National University of Ireland, Galway

in fulfillment of the requirement for the degree of Doctor of Philosophy,

in the School of Geography, Archaeology and Irish Studies

Submission date: September 2019

Table of contents

Declarations

This thesis or any part thereof, has not been, or is not currently being submitted for any degree at any other university.

Yuting Meng

The work reported herein is as a result of my own investigations, except where acknowledged and referenced.

Yuting Meng

Abstract

As geographical information system (GIS) and spatial analysis techniques have been widely used in soil contamination, the methodologies applied in urbanized area and submillimetre profiles widened their utility. In this study, the spatial distribution patterns of phosphorus (P), aluminium (Al), calcium (Ca) and lead (Pb) in urban topsoils and labile P in lake sediments have been investigated using GIS-based techniques. The valuable results would provide information for soil management.

The P is a necessary and limiting nutrient element for plants. The excessive P is becoming a worldwide environmental concern because it could cause water eutrophication, then threaten natural ecosystems and even human life. The hot spot map and results of analysis of variance (ANOVA) showed a mixed natural and anthropogenic control on P in the Greater London Authority (GLA) area based on analysis of 6467 topsoil samples. In the aspect of natural influences, alluvium and river terrace deposits had elevated P content. The hot spot map clearly demonstrated that the hot spots in the lower Thames Estuary disappeared suddenly, indicating the influences of the tidal dilution on the P-enriched soils. The high values clusters of hot spots in the city center and the built-up areas suggested that soil P content was heavily affected by human activities, which had also been revealed by the significant difference between the P concentration in urbanized area and that in non-urbanized area.

The GIS techniques were used to visualize the labile P distribution in sediments of Lake Taihu at a submillimetre scale, showing low labile P in winter and high labile P in summer. The significant spatial autocorrelation in sediment profiles was revealed by the high values of Moran's I. The distinct submillimeter-scale spatial patterns with seasonal changes were exhibited by the semivariograms. The impact of temperature on the mobility and spatial distribution of labile P in sediments was revealed by the apparently different levels of spatial variation during summer and winter.

Except for P, the spatial distribution of chemical elements in urban topsoils can be changed with different degrees. Three metals, Al, Ca and Pb in urban topsoils of the GLA area, displayed different levels of geogenic and anthropogenic influences, revealed by ANOVA and GIS-based spatial analysis. The spatial distribution of Pb was remarkably changed by human activities; The inert element Al could still be able to reserve its natural spatial distribution in the highly urbanized London and Ca exhibited the hybrid spatial distribution affected by both natural factors and human activities.

Since Pb is often affected by human activities and it is a toxic heavy metal that is detrimental to human health as well as social behavior, we attempted to prepare the investigation of the relationship between Pb and socioeconomic or healthy status in the GLA area in the future. The first challenge is the spatially misaligned data when linking point Pb concentration to the socioeconomic or healthy data of the ward. This study has highlighted the point-to-area transformation issue on the basis of review and comparison of various methodologies. The advantages and disadvantages of different methods have been pointed out. The method of median block inverse distance weighted (IDW) interpolation was recommended for our further investigations.

Keywords: Phosphorus; GIS; London; Moran's I, local Moran's I; Semivariogram; Lake Taihu; Lead; point-to-area data transformation

Acknowledgments

Primarily, I would like to express my sincere gratitude to China Scholarship Council (No. 201504910793) and my supervisor Dr. Chaosheng Zhang, for giving me the opportunity to start my Ph.D. studies in the department of Geography, National University of Ireland, Galway. I would like to thank Dr. Chaosheng Zhang for continuous support and guidance in carrying out research, writing papers, cooperating with other professionals. Special thanks to his wife Mrs. Dongxia Zhang and their children who cared my life and made me feel at home.

I would like to thank Dr. Mark Cave, for his thorough reviewing my manuscript and providing insightful suggestions.

I would like to thank Prof. Shiming Ding, Dr. Aaron Potito and Dr. Karen Daly as the members of my graduate research committee for their valuable comments on the progress of my research project.

I would like to thank Dr. Siubhán Comer for her technical assistance and caring my life.

I would like to thank all the friends and colleagues in Geography department, who made my last four years full of warmth and joy. Special thanks to Yumin Yuan, with whom I spent four years studying, working together. She is helpful and considerate. Special thanks to Elaine Williams, who is always friendly and kind.

At the end, I would like to thank my loving family. Words cannot express how grateful I am to my parents, parents-in-law, husband Yicheng Ding and son Wenli Ding. Their constant support and company make me feeling beloved and would encourage me to climbing a higher mountain.

Publications

This thesis comprises 4 publications as the first author.

1) **Meng, Y.,** Cave, M., Zhang, C., 2018. Spatial distribution patterns of phosphorus in topsoils of Greater London Authority area and their natural and anthropogenic factors. Applied Geochemistry 88, 213-220.

2) **Meng, Y.,** Ding, S., Gong, M., Chen, M., Wang, Y., Fan, X., Shi, L., Zhang, C., 2018. Submillimeter-scale heterogeneity of labile phosphorus in sediments characterized by diffusive gradients in thin films and spatial analysis. Chemosphere 194, 614-621.

3) **Meng, Y.,** Cave, M., Zhang, C., 2019. Comparison of methods for addressing the point-to-area data transformation to make data suitable for environmental, health and socio-economic studies. Science of the Total Environment 689, 797-807.

4) **Meng, Y.,** Cave, M., Zhang, C., 2020. Identifying geogenic and anthropogenic controls on different spatial distribution patterns of aluminium, calcium and lead in urban topsoil of Greater London Authority area. Chemosphere 238, 124541.

List of Abbreviations

1D	One-dimensional
2D	Two-dimensional
Al	Aluminium
ANOVA	Analysis of variance
BGS	British geological survey
Ca	Calcium
Cd	Cadmium
Cr	Chromium
Cu	Copper
DGT	Diffusive gradients in thin films
EU	European union
Fe	Iron
GIS	Geographical information system
GLA	Greater London authority
IDW	Inverse distance weighted
LISA	Local indicators of spatial association/autocorrelation
Mn	Manganese
N	Nitrogen
Ni	Nickel
P	Phosphorus
Pb	Lead
PM	Particulate matter
PMs	Parent materials

PVDF	Polyvinylidene difluoride
SWI	Sediment-water interface
UK	United Kingdom
XRF	X-ray fluorescence
Zn	Zinc

Chapter 1
Introduction

1.1 Background

Geographic information system (GIS) mapping and spatial analysis have been widely used to reveal spatial distributions and structures of soil properties (Ashaiekh et al., 2019). The GIS applications provide invaluable information for risk assessment and soil management (Meng et al., 2018). The GIS techniques are powerful for determination of spatial distribution and behavior of geochemical elements in urban areas (Mihailović et al., 2015; Kim and Choi, 2017). Parchomenko and Borsky (2018) applied the Getis-Ord Gi* statistic to identify a phosphorus (P) hot spot where was oversupplied by manure and recommended a spatially-targeted regulation on manure application. Kumar et al. (2018) used Moran's I to choose the threshold distance for the Getis-Ord Gi* statistic, then they assessed the risk of elevated uranium in groundwater depending on the hot spot analysis. Since chemical elements in urban soils have been intensively altered by human activities, different elements exhibit different levels of influences. Identifying the differences helps to discover accumulation locations and potential sources, which is useful to better manage and control metal accumulation or contamination in soils. In this study, the GIS techniques, geostatistics, spatial analysis were applied to investigate different chemical elements (P, aluminium (Al), calcium (Ca) and lead (Pb)) at different spatial scales (regional scale, submillimeter scale, point-to-area scale).

The P is an indispensable element as well as a limiting nutrient in terrestrial and aquatic ecosystems. Phosphodiester is a component of nucleic acids and adenosine triphosphate is an energy transport molecule in a cell. In the natural (pre-human) P cycle, the only significant source is the chemical weathering from apatite minerals (Wu et al., 2016). Human activities have dramatically influenced P cycle in a variety of forms. Among all, the application of P-fertilizer in agriculture is the main source of soil P input. Between 1850 and 2000, the global farmland received approximately 550 Mt P (Smil, 2000). The other markable sources include deforestation, waste and sewage. In order to meet the demands of food consumption and urbanization, more and more forests are cut. Deforestation increases P losses via soil erosion and runoff (Maranguit et al., 2017). Huge amounts of animal manure are discharged to the environment, causing serious environmental and economic

problems due to P losses (Jia et al., 2015). Although P in sewage is largely controlled nowadays, the sludge produced by sedimentation of sewage still retains more than 50% of P after standard treatment (Orhon, 2015). The P from diffuse sources through soil leaching and from point sources, such as discharge of sewage, are often responsible for the eutrophication of water bodies (Némery et al., 2005) (Fig. 1). In spite of low solubility of phosphates and their rapid transformation to insoluble forms, P can convert shallow freshwater bodies from oligotrophic to hypereutrophic. The concerns bringing from eutrophication range from unacceptable taste of drinking water to toxic matters (e.g. cyanotoxins) released by algal blooms (Oliver et al., 2019). The best way to minimize the anthropogenic impacts on P is to reduce its input. However, it is not fit for the high-consumption societies of today. Improving the efficiency of P utilization and controlling the discharge of P are alternative choices. Setting more effective policies and management methods requires a better understanding of the location and magnitude of P in soil. For instance, building tree buffer zones and preventing uncontrolled feedlot waste can moderate P runoff, which can be located using GIS mapping and spatial statistical methods. Besides soil P leaching, internal P release from sediments are another vital cause of eutrophication. A large number of lakes in China are encountering the challenge of eutrophication. Two-dimensional (2D) high-resolution profiling of P measured by diffusive gradients in thin films (DGT) at the sediment-water interface (SWI) has been studied extensively in recent years (Chen et al., 2019a; Chen et al., 2019b). The distribution of DGT-labile P improves the understanding of the biogeochemistry behavior of P. However, classical statistic description often turns the 2D spatial distribution profile to one-dimensional (1D) graph, using the mean values and standard deviation of each depth, or even several millimeters. This would lead to loss of information. Recently, researchers have utilized geostatistical theories to quantitatively study spatial correlation (Onyejekwe et al., 2016; Zhao et al., 2018). The spatial variation of sub-millimeter scale DGT-labile P concentration has not been quantified using geostatistics yet, which is worth to be investigated.

Three metal elements, Al, Ca and Pb, were selected to compare the extent of geogenic control and anthropogenic influences in the GLA area. It was found that Pb is vulnerable to anthropogenic influences and Al is resistant to anthropogenic influences. It is well established that urban soils

often contain elevated Pb concentrations (Spliethoff et al., 2016). Exposure to Pb has potential hazards to human health. Studies have shown that Pb exposure would affect nervous, cardiovascular, immunological systems, leading to intelligence quotient deficits (Taylor et al., 2017), hearing loss (Carlson and Neitzel, 2018) and asthma (Farkhondeh et al., 2015). In addition, Pb has negative effects on socioeconomic status and psychophysiological behaviors (Gump et al., 2009).

Furthermore, a challenge of different supports of data sets emerged when we prepared to carry out a research about the association between soil Pb and socioeconomic or healthy status. Geochemical data are often collected as point data, for example the geochemistry concentration in soils; On the other hand, the health data (e.g. mortality rate, cancer morbidity rate) are usually aggregated in administration areal unit level. The problem of point-to-area scale transformation is interesting to be reviewed and investigated.

1.2 Policies on soil contamination of European Union, United Kingdom and China

Soil is an essential and non-renewable natural resource and offers primary functions, including food production, hydrological management and maintenance for sustainable biological diversity and terrestrial ecosystems. The Thematic Strategy for Soil Protection was adopted with the objective to protect soils, identify the soil problem and remediate contaminated soil across the European Union (EU). Member States would support research projects, such as in the areas of soil nutrients recycling, consolidate soil monitoring and investigations, operate with contaminated sites and further integrate soil and soil protection aspects in different policies (EC, 2006). The Environmental Protection Agency of Ireland is the authority to carry out EU regulations that cover soil protection. The Environmental Protection Agency encourages soil research in order to precisely assess soil states and to supply information and guidance to policy and decision makers (EPA, 2014). Department for Environment, Food and Rural Affairs of the United Kingdom (UK)

has supported a large amount of research to better understand the risk of soil pollution for the purpose of safeguarding soils (Defra, 2009). The Ministry of Ecology and Environment of People's Republic of China has made legislation on protection and remediation of soil pollution and black odors water bodies (Xu et al., 2017).

1.3 Research objectives

The spatial distribution patterns of P, Al, Ca and Pb in London topsoils and P in Lake Taihu sediments have been modified by human activities in different ways. The main objectives are to investigate the spatial distribution patterns and their influencing factors using GIS techniques and spatial analysis.

The specific objectives for spatial variation of regional-scale P distribution in topsoils of Greater London Authority (GLA) area include:

1) to reveal the spatial patterns of P using GIS mapping tool;
2) to identify the locations of P accumulation and depletion using local Moran's I;
3) to explore the natural and anthropogenic factors associated with the spatial patterns.

The specific objectives for spatial variation of submillimeter-scale DGT-labile P distribution in sediments of Lake Taihu include:

1) to visualize the monthly distribution of labile P in sediments using GIS techniques so as to reveal the seasonal differences of distribution;
2) to quantify spatial autocorrelation of labile P using spatial analysis methods, including semivariogram and Moran's I, in order to determine its spatial structure characteristics at the submillimeter scale.

The specific objectives for identification of geogenic and anthropogenic controls on different spatial distribution patterns of Al, Ca and Pb include:

1) to visualize spatial distribution of Al, Ca and Pb in topsoil of the GLA area using spatial interpolation;
2) to identify the different levels of influences of natural and anthropogenic factors on their spatial distributions using analysis of variation (ANOVA) and GIS-based spatial analyses;
3) to explore the natural and anthropogenic factors associated with the spatial patterns.

The specific objectives for comparison of methods for addressing the point-to-area data transformation to make data suitable for environmental, health and socio-economic studies include:

1) to review recently used methods of linking point data to area data;
2) to compare the popular methods of estimating concentration of soil Pb and to select one "optimal" methodology for the further study.

1.4 Structure of thesis

This thesis comprises 5 chapters:

Chapter 1 is the introduction, presenting the background, main objectives and research structure.

Chapter 2 reviews the literatures of following aspects:

1) The application of GIS techniques and spatial analysis methods;
2) The measurements, influencing factors of spatial P variation in urban soils and lake sediments;
3) The general spatial information of Al, Ca and Pb, the hazard of Pb contamination

4) The causes and resolutions of point-to-area scale transformation problem.

Chapter 3 demonstrates the materials and methodologies used in the research.

Chapter 4 consists of four published papers with their summaries and personal dedication descriptions.

Chapter 5 is a discussion chapter which highlights how the four papers:

1) relate to each other;

2) contribute to the study's stated goals and link to the literature reviewed;

3) advance the field of soil contamination, and the utility of GIS-based techniques.

Finally, Chapter 6 concludes the four papers, and recommends policy relevant strategy and puts forward future research.

Reference

Ashaiekh, M. A., Eltayeb, M. A. H., Ali, A. H., Ebrahim, A. M., Salih, I., Idris, A. M., 2019. Spatial distribution of total and bioavailable heavy metal contents in soil from agricultural, residential, and industrial areas in Sudan. Toxin Reviews 38(2), 93-105.

Carlson, K., Neitzel, R., 2018. Hearing loss, lead (Pb) exposure, and noise: a sound approach to ototoxicity exploration. J. Toxicol. Env. Health, Part B 21(5), 335-355.

Chen, M., Ding, S., Wu, Y., Fan, X., Jin, Z., Tsang, D. C. W., Wang, Y., Zhang, C., 2019a. Phosphorus mobilization in lake sediments: Experimental evidence of strong control by iron and negligible influences of manganese redox reactions. Environ. Pollut. 246, 472-481.

Chen, Q., Chen, J., Wang, J., Guo, J., Jin, Z., Yu, P., Ma, Z., 2019b. In situ, high-resolution evidence of phosphorus release from sediments controlled by the reductive dissolution of iron-bound phosphorus in a deep reservoir, southwestern China. Sci. Total Environ. 666, 39-45.

Defra (Department for Environment, Food and Rural Affairs), 2009. Safeguarding our Soils: a strategy for England. Nobel House, 17 Smith Square, London SW1P 3JR. Available online at: http://defraweb/environment/land/soil/index.htm (Accessed on 2 Jul.19).

EC (European Commission), 2006. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions. Thematic strategy for soil protection COM(2006) 231, European Union, Brussels, Belgium.

EPA (Environmental Protection Agency), 2014. Irish soil information system project. EPA and Teagasc, Wexford, Ireland.

Farkhondeh, T., Samarghandian, S., Sadighara, P., 2015. Lead exposure and asthma: an overview of observational and experimental studies. Toxin Rev. 34(1), 6-10.

Gump, B. B., Reihman, J., Stewart, P., Lonky, E., Granger, D. A., Matthews, K. A., 2009. Blood lead (Pb) levels: further evidence for an environmental mechanism explaining the association

between socioeconomic status and psychophysiological dysregulation in children. Health Psychol. 28(5), 614-620.

Jia, W., Yan., Z., Chadwick, D. R., Kang, L., Duan, Z., Bai, Z., Chen, Q., 2015. Integrating soil testing phosphorus into environmentally based manure management in peri-urban regions: A case study in the Beijing area. Agr. Ecosyst. Environ. 209, 47-59.

Kim, S., Choi, Y., 2017. Assessing statistically significant heavy-metal concentrations in abandoned mine areas via hot spot analysis of portable XRF data. Int. J. Environ. Res. Public Health 14(6), 654. https://doi.org/10.3390/ijerph14060654

Kumar, D., Singh, A., Jha, R. K., Sahoo, S.K., Jha, V., 2018. Using spatial statistics to identify the uranium hotspot in groundwater in the mid-eastern Gangetic plain, India. Environ. Earth Sci. 77, 702. https://doi.org/10.1007/s12665-018-7889-1

Maranguit, D., Guillaume, T., Kuzyakov, Y., 2017. Land-use change affects phosphorus fractions in highly weathered tropical soils. Catena 149, 385-393.

Meng, Y., Cave, M., Zhang, C., 2018. Spatial distribution patterns of phosphorus in topsoils of Greater London Authority area and their natural and anthropogenic factors. Applied Geochemistry 88, 213-220.

Mihailović, A., Budiinski-Petković, L., Popov, S., Ninkov, J., Vasin, J., Ralević, N.M., Vasić, M. V., 2015. Spatial distribution of metals in urban soil of Novisad, Serbia: GIS based approach. J. Geochem. Explor. 150, 104-114.

Némery, J., Garnier, J., Morel, C., 2005. Phosphorus budget in the Marne Watershed (France): urban vs. diffuse sources, dissolved vs. particulate forms. Biogeochemistry 72, 35-66.

Oliver, S., Corburn, J., Ribeiro, H., 2019. Challenges regarding water quality of eutrophic reservoirs in urban landscapes: a mapping literature review. Int. J. Environ. Res. Public Health 16, 40.

Onyejekwe, S., Kang, X., Ge, L., 2016. Evaluation of scale of fluctuation of geotechnical parameters by autocorrelation function and semivariogram function. Eng. Geol. 214, 43-49.

Orhon, D., 2015. Evolution of activated sludge process: the first 50 years. J. Chem. Technol. Biotechnol. 90, 608-640.

Parchomenko, A., Borsky, S., 2018. Identifying phosphorus hot spots: a spatial analysis of the phosphorus balance as a result of manure application. J. Environ. Manage. 214, 137-148.

Smil, V., 2000. Phosphorus in the environment: natural flows and human interferences. Annu. Rev. Energy Environ. 25, 53-88.

Spliethoff, H. M., Mitchell, R. G., Shayler, M., Marquez-Bravo, L. G., Russell-Anelli, J., Ferenz, G., McBride, M., 2016. Estimated lead (Pb) exposures for a population of urban community gardeners. Environ. Geochem. Hlth. 38(4), 955-971.

Taylor, C. M., Kordas, K., Golding, J., Emond, A. M., 2017. Effects of two-level prenatal lead exposure on child IQ at 4 and 8 years in UK birth cohort study. Neurotoxicology 62, 162-169.

Wu, H., Zhang, Y., Yuan, Z., Gao, L., 2016. A review of phosphorus management through the food system: identifying the roadmap to ecological agriculture. J. Clean. Prod. 114, 45-54.

Xu, C., Yang, W., Zhu, L., Juhasz, A. L., Ma, L. Q., Wang, J., Lin, A., 2017. Remediation of polluted soil in China: policy and technology bottlenecks. Environ. Sci. Technol. 51, 14027-14029.

Zhao, L., Zhang, S., Huang, D., Zuo, S., Li, D., 2018. Quantitative characterization of joint roughness based on semivariogram parameters. Int. J. Rock Mech. Min. 109, 1-8.

Chapter 2
Literature Review

2.1 Overview

This chapter reviews the literatures of the significance of GIS techniques and spatial analysis methods and its applications. As the powerful GIS-based analysis has been widely used, it can be applied to investigate spatial distribution patterns of chosen chemical elements (P, Al, Ca and Pb) with different spatial scales in urban soils and in lake sediments. This chapter also refers the measurements, influencing factors of spatial P variation in urban soils and lake sediments, the general spatial information of Al, Ca and Pb, the hazard of Pb contamination, and the causes and resolutions of point-to-area scale transformation problem.

2.2 Introduction

GIS technology is not only used to produce maps, but also to help researchers to identify the spatial distribution patterns and find out new phenomenon. The new findings would inspire a deeper understanding of the mechanisms. In the past decades, enhanced visualization tools and spatial analyses have been widely applied in interdisciplinary researches, especially in environment-related and health-related fields. In this study, the GIS-based analysis is used to explore the spatial distribution patterns of P in extremely urbanized area and in microscale profile.

Since extensive urbanization, large areas of agriculture, grassland or forest soils have been changed into urban soils. The geochemistry and other characteristics of urban soils are tremendously altered by human activities, such as constructing buildings, transforming inputs and outputs of materials through the city (Wei et al., 2019). The P plays a significant role in urban soil and water quality, closely related with environment as well as residents' health. However, elevated P has been found in cities due to greater input than output (Li et al., 2011). The excessive P would potentially transport into aquatic ecosystems and lead to the

eutrophication in water body (Addiscott and Thomas, 2000). In the UK, Pretty et al. (2002) preliminarily assessed that the economic loss caused by freshwater eutrophication was 75-114 million pounds annually. It is believed that sewage treatment is one of the point sources of P, while the soil P accumulation is a sort of diffuse source as well (Granger et al., 2017). It is, therefore, of importance to monitor the locations of soil P accumulation.

The soil P transports from accumulated-P soils to water bodies through various water pathways, for instance, leaching to groundwater, and surface run-off to stream (McGinley et al., 2016). Except for the external P input, bottom sediments of lakes act as a sink of P as well as a source of P (Chen et al., 2018). When the balance between sediment and water phase is broken, P would exchange through SWI and it is frequently happened in the uppermost layer of bottom sediment (Biswas et al., 2017). A recent study showed that the internal release of P from sediment, rather than external P loading, determined the P limitation for harmful algal blooms in hyper-eutrophic Lake (Wu et al., 2017). Thus, it is necessary to investigate the spatial variation of P in water sediments.

Except for P, metal elements in urban soil have been modified by human activities in different levels. It is of great interest that three chosen typical metal elements, Al, Ca and Pb clearly demonstrate the different levels affected by human activities in an extremely urbanized city. Understanding how much human activities influence the spatial variation is important for soil management.

Moreover, if researchers attempted to investigate the relationship of urban soil geochemical with environment, health or socio-economy, they would encounter with incompatible data of different spatial scales. Addressing this spatially incompatible data problem is an important step which will influence the further investigation. In this chapter, this problem is reviewed.

2.3 The GIS technology

The GIS technology is commonly used in the study of soil property at the beginning. Xu et al. (2019) revealed the "hidden" relationship and pattern between total organic carbon and pH at the European scale using hot spot analysis. Fu et al. (2010) investigated the spatial variation of P in a grassland of Ireland using local Moran's I. Meng et al. (2018(b)) identified the natural and anthropogenic factors that influence the P spatial distribution pattern in the GLA area using local Moran's I.

The GIS technology has been applied in environmental epidemiology study, in order to better understand the changes of the Earth's surface by natural, social and economic processes that affect public health on local, regional and global scales (Cromley and McLafferty, 2002). The cases range from basic disease mapping to spatial analysis and spatial modeling. The GIS techniques can provide powerful spatial analysis tools to investigate spatial characteristics of disease and its environmental influence factors (Rapant et al., 2015). Zhang and Tripathi (2018) used IDW interpolation, global Moran's I statistics, Getis-Ord G* statistics and local indicators of spatial association/autocorrelation (LISA) to examine the correlation between lung cancer and fine particulate matter ($PM_{2.5}$).

The GIS also can help to optimize the location for business or facilities. Xue et al. (2017) found the optimal location of urban mining facilities in China with the assistance of GIS. Woo et al. (2018) optimized the location of biomass energy facilities using the mixed method of multi-criteria analysis and GIS.

The GIS provides spatial-temporal tools for investigating the change extent as well as identifying the area with high risk. García-Ayllón (2017) applied GIS spatial-temporal analysis to assess anthropogenic influences on coastal ecological systems, such as land use, while Singh

and Kasana (2017) performed GIS-based spatial-temporal investigation of groundwater level fluctuation. Vemulapalli et al. (2017) identified distinct spatial and temporal patterns of locations with high risk for aging-involved roadway accidents.

The semivariogram and kriging method, initially used for mining industry, is now widely employed in environment sciences (Oliver and Webster, 2015). The geostatistics method is also applied to the material (Niemietz et al., 2010) and submillimetre scale (Meng et al., 2018(a)) researches. Niemietz et al. (2010) carried out variogram analysis to analyze physical properties of semiconductors.

2.4 The GIS application of P in soils

Since GIS-based assessment and management of P have been widely applied in catchment and agricultural land (Gupta et al., 2012; Fu et al., 2010; Fu et al., 2011), the P in urban soils has not been fully investigated. In 2016, several eutrophic phenomena have been observed during London field trip (Fig. 2.1). The spatial variability in urban soils is complicated due to human disturbance which needs GIS technology to identify the spatial distribution patterns. The spatial distribution patterns will provide valuable information for P management.

Figure 2.1: (a, b, c) One section of Lee Navigation; (d) The stream near the Deephams Sewage Work; (e) Deephams Sewage Work; (f) Trunk Sewers on Meridian Way; (g, h, i) Stanwell Village.

2.4.1 Spatial variability of P in soils

Under natural condition, the spatial variability of P is influenced by complicated factors, such as parent materials (PMs), topography, climate, hydrology, vegetation types, soil texture and biotic processes (Bennett et al., 2004; Guan et al., 2017). In mountain area, PMs play a determining role in the soil P cycling (Mage and Porder, 2013). In humid tropical forest, soil moisture affects the available P with a positive relationship (Wood et al., 2016). In agricultural areas, topographic index is related with the extractable P under the condition of low external P input (Wilson et al., 2016). In one forest of Beijing, vegetation types and elevation have strong impacts on soil P (Zhang et al., 2010). On the other hand, P accumulation and storage are altered by land use, fertilizer use and manure applications (Hu et al., 2011). The distribution of P in farms can be significantly modified by land uses, such as cropland, grassland and mountain pasture (Roger et al., 2014; Xu et al., 2014). Fertilizing increases soil P dramatically (Gao et al., 2011) and manure results in high P surpluses (Jia et al., 2015).

Soil properties exhibit variance and heterogeneity at different spatial scales, ranging from submillimeter to kilometer (Yost et al., 1982; Hu et al., 2003; Meng et al., 2018(a)). In many circumstances, spatial variation is not random but has autocorrelation in which variability increases as distance increases between points (Fu et al., 2011). The precision of P management depending on the identification of P variability can significantly reduce the risk of P accumulation. In practice, farmers focus on the variability within field; while environmental planners and policy makers are interested in the regional variability. Previous studies paid attention on agriculture land and rural area, however P in urban soil has not been fully studied. Urban system changes the input and output of P, thus soils in several cities such as Nanjing, Hangzhou, Hefei of China, have been enriched with P (Zhang et al., 2001; Zhang, 2004; Li et al., 2010). Urban soil is closely related to health due to huge populations living in urban areas. It is therefore vital to investigate the magnitude and location of P accumulation and storage. The research sets the foundation for improving P management and preventing P from transporting into aquatic systems.

2.4.2 Anthropogenic factors for P variability

2.4.2.1 Fertilizer

To meet the rapidly increasing demand for food production, P input has been transformed from locally available organic matter like manure and human excreta to inorganic P fertilizer derived from distant phosphate-rich rock since the mid-to-late 19th century. It is predicted that the P demand would continue to grow until 2030 and the P scarcity would threaten the food supply on global level (Cordell et al., 2009). The global agricultural soils received around 550 Mt P from 1850 to 2000 (Smil, 2000). In the UK, P losses from agricultural land have contributed to eutrophication, although P surpluses have decreased by about 50% due to the reduction of P fertilizer application since the last decades (Ulén, et al., 2007). Except for agricultural land, fertilizer is also widely used in urban domestic gardens which is often neglected but is significant for the urban soils. Gardens in Flemish have higher P content due to the usage of 0.07 kg fertilizer per m^2 (Dewaelheyns et al., 2013). London has a large area of garden of 37,900 hectares, accounting for approximately 24% of the GLA area (Smith et al., 2011). Thus, the fertilizer application in gardens plays an important role in urban soil. In China, the consumption of P fertilizer has boosted about 100-fold during the period of 1960 and 2010 (Li et al., 2015). Despite of the improved efficiency, there are still plenty of P exported from river basins to the coastal areas (Liu et al., 2018(a)).

2.4.2.2 Reuse of organic wastes (manure)

Because of scarcity and increasing cost of exploitable phosphate rock, over-abundance of crop residues and animal manure are recycled to meet the need of P (Sarvajayakesavalu et al., 2018). Government encourage to reuse livestock manure because it remains valuable plant nutrients and organic matter (Grande et al., 2005). There are around 2 Gt dry matter produced by fresh

manure per year worldwide. In Europe and Asia, agricultural fields received 5-10 kg P per hectare annually which is contributed by animal manure (Smil, 2000). However, the application of manure is spatially uneven. Some areas have manure P deficits due to unavailability of local livestock and the bulky texture of manure, while some areas have saturated or even surplus P (Bateman et al., 2011). Although the addition of manure is beneficial to agricultural sustainability, repeated manure application for achieving the balance of nitrogen (N) and P would cause high P accumulation (Bundy and Sturgul, 2001). In China, the huge amounts of livestock and poultry manure disposed to the environment have brought up the P pollution (Yuan et al., 2019). The high manure P surpluses in Beijing, China resulted from per-urban arable land have become a big concern regarding the environment. Balanced manure P management is urgently required to minimize the P uneven distribution and potential risk of runoff.

2.4.2.3 Sewage

Organic waste and detergent component (sodium triphosphate) are important sources of P in domestic wastewater (Halliwell et al., 2001). Wastewater is centrally treated in large cities nowadays, which is a more economic and efficient way for disposal of human waste to water. This step transferred multiple small diffuse sources of P to large point sources. Across the United States of America (USA), nearly 0.06 Tg P transported to riverine as sewage and detergent in 2012 (Metson et al., 2017). In the UK, with the great efforts made by sewage treatment works, the mean total reactive P concentration in river waters has decreased by 60% since a peak in 1985 (Civan et al., 2018). Although P has been removed partially by treatment works prior to discharge into receiving waters, downstream of the receiving water still contains high concentrations of all species of P, particularly soluble reactive P (Millier and Hooda, 2011). On the other hand, millions of tons of sewage sludge are generated globally, which remains high P content (1-5%) (Wang et al., 2018.). Utilization in agriculture, landscaping, incineration and landfilling are major methods for sludge treatment (Kelessidis and Stasinakis, 2012). Sewage sludge is not suitable to reuse in agriculture directly owning to harmful heavy metal

residues. Thus, regenerating P from sewage sludge is an environmental-friendly alternative (Wollmann et al., 2018). For example, sewage sludge ash can be made into P fertilizer (Herzel et al., 2016). Inappropriate disposal of sewage and its derived substance would likely export excessive P into the aqueous environment. Most eutrophication is attributable to the sewage emitting from boosting urbanized human population (Reynolds and Davies, 2001).

2.4.3 The P management

It is urgent to reduce anthropogenic P burdens as it is already an awareness that unsustainable use of P has led to worldwide eutrophication. GIS techniques can help in P management which has been mentioned in the following contents. The P management can be divided into three general categories: cutting down the P input; mitigating P loss to water body; enhancing the recovery and reuse of P waste.

Cutting down the P input:

1) Improve efficiency of P fertilizer in farming system. Spatial variation of nutrient in agricultural land has been recognized, especially by GIS. Thus, precise application of fertilizer is a way to improve financial and environmental outcomes (Fu et al., 2010). On the other hand, optimizing input rate of fertilizer based on assessing the efficiency by several methods, such as P isotope track, P Balance Efficiency is also an approach (Iho and Laukkanen, 2012; Weaver and Wong, 2011).

2) Limit the number of livestock and the amount of manure produced in problem areas (Smil, 2000).

3) Improve farmers' awareness of the negative consequences of P fertilizer. Some surveys have found that farmers over-applied fertilizer in order to increase yields (Li et al., 2016).

4) Restrict the use of P-containing detergents. In 2012, detergent was still one of the major sources of P water pollution in the USA (Metson et al., 2017).

5) Increase the performance of fertilizer by technology, such as addition of P-solubilizing bacteria in fertilizer (Abbasi et al., 2015).

6) Encourage the change of diet type of people by reducing meat and dairy consumption. Given the fact that more than 60% of fertilizer is used on cereals of which more than 40% is used as livestock feed (Smil, 2000).

Mitigating P loss to water body:

1) Increase the capability of removal of P from sewage and industrial. Developing techniques of biomass engineering and sewage treatment works is a vital step to minimize the transportation of P into aqueous systems (Chan et al., 2017).

2) Decrease the speed of deforestation. Abundant organic matter in forest soils is helpful in reducing water and nutrients runoff (Ouyang et al., 2015).

3) Establish treed buffer zones around intensively fertilized fields, high P concentration areas or along riverine. GIS mapping and hot spot analysis could help to identify the areas with high P pollution risk (Meng et al., 2018(b)).

4) Slow down the flow rate during rainfall events with high intensities and short durations. Li et al. (2013(a)) pointed out that flow rate is the primary influence factor on P release in urban soils.

Enhancing the recovery and reuse of P waste:

1) Balance P among the arable land, grassland and livestock sectors. The efficiency of manure collection, storage and field application are crucial in reuse of P waste. The closer linkage of these sectors would reduce P losses (Li et al., 2016).

2) Improve manure to be more suitable for farm application by techniques. For example, fungal biomass treatment can increase the N-to-P ratio which can better satisfy crop growth and help to reduce P pollution (He et al., 2019). Extracted P from manure by acid

wash can be used as fertilizer and the manure after this treatment is environmentally safe for land application (Szögi, et al., 2015).

3) Recover P from sewage. Combined with population growth and urbanization development, the potential of recovery P from sewage is huge (Wu et al., 2016).

2.5 The GIS application of labile P in lake sediments

Internal labile P in lake sediments plays an important role in water eutrophication. Labile P can be available to plants or microorganisms, which includes the sum of inorganic P and organic P extracted in resin and bicarbonate (Costa et al., 2016). It can be captured in 2D submillimeter-scale profiles. However, it is often described as 1D. The 2D spatial information obtained by GIS can help in decision making and support for better management of the water resources.

2.5.1 Spatial variability of labile P in lake sediments

After an excess of P moves to lakes, the sediments perform not only as a sink of P by forming mineral-P complexes, but also as a source to release P back to overlaying water column (Xu et al., 2013). Therefore, internal loading of P can be sufficient to maintain waterbody in a eutrophic state for a long period of time, even when the external anthropogenic input has already substantially declined (Wang et al., 2011). Although porewater dissolved reactive P in sediment is increasing from top to lower depth, P flux is more active in the top few centimeters of the sediment. It is because that the anaerobic condition of uppermost sediment promotes a large amount of P fluxes (Moore et al., 1998). As a result, a debate on the effects of dredging on the control of P is ongoing. Some studies pointed out that dredging is an effective way to reduce the internal P loading (Liu et al., 2016(a); Yu et al., 2017); others reported as an opposite conclusion (Pu et al., 2000). The spatial variation of P does not only exist in different locations

(Spears et al., 2006; Yao et al., 2016), but also presents in different depths, even in submillimeter scale in sediments (Meng et al., 2018(a); Yao et al., 2016). Quantification of spatial variability is essential for investigating the relationship between water sediment properties and environmental factors. Management of eutrophication depends on the determination of spatial horizontal and vertical variation as well as the mechanism behind it.

2.5.2 Measurements of labile P in lake sediment

The assessment of total P in sediment gives essential information about its enrichment, while the measurement of bioavailability of P is vital for measuring its mobilization (Barik et al., 2019). Sequential extraction using different solution is a traditional method for analysis of P fraction in water and sediments (Chen et al., 2016(a); Moore et al., 1998; Tang et al., 2018). The measurement at high spatial resolution and in situ is required at SWI because the chemical varies within short distance in the surface of sediment (Ding et al., 2011). Several techniques have been developed recently, such as Rhizon, diffusive equilibration in thin films (DET) and DGT (Ding et al., 2011; Xu et al., 2012). Precise data from such measurements are necessary in order to improve our knowledge of biogeochemical processes.

The DGT technology was invented by Bill Davison and Hao Zhang in 1993 which is based on the fundamental theory of Fick's first law of diffusion. It is a novel approach for measuring the labile forms of chemical elements, for instance P, sulfur (S) and arsenic (As), in water, sediment and soil (Davison and Zhang, 1994). The DGT has the following merits: in situ measurement; labile fractionation; kinetic and thermodynamic constant; high spatial resolution; 2D concentration images (Zhang et al., 2014). The device comprises three layers, including binding gel, diffusive gel and filter membrane (Fig. 2.2). The binding gel contains binding agents to accumulate ions. It is the establishment of a constant concentration gradient in the diffusive layer that forms the basis for measuring chemical element concentration in solution quantitatively. The labile ions diffuse through a filter membrane and a diffusive gel and are

then bound on the binding gel. The concentration of labile P can be calculated using Equation (2.1) (Davison and Zhang, 1994):

$$C_{DGT} = \frac{M\Delta g}{DAt} \quad (2.1)$$

Where Δg (cm) represents the thickness of the diffusive layer, D ($cm^2\ s^{-1}$) refers to the diffusive coefficient of P in the diffusive layer, t (s) is the deployment time, A (cm^2) represents the exposure area of the gel, and M (μg) is the accumulated mass over the deployment time (e.g., 24 hours).

In reality, the 2D concentration images obtained by the DGT technology are often changed to 1D vertical distributions using conventional statistics (Xing et al., 2018; Yao et al., 2016). This change would lose a lot of valuable spatial information. There is a gap between the 2D image information and conventional statistical analysis. The GIS techniques and geostatistics are proposed to handle with this issue.

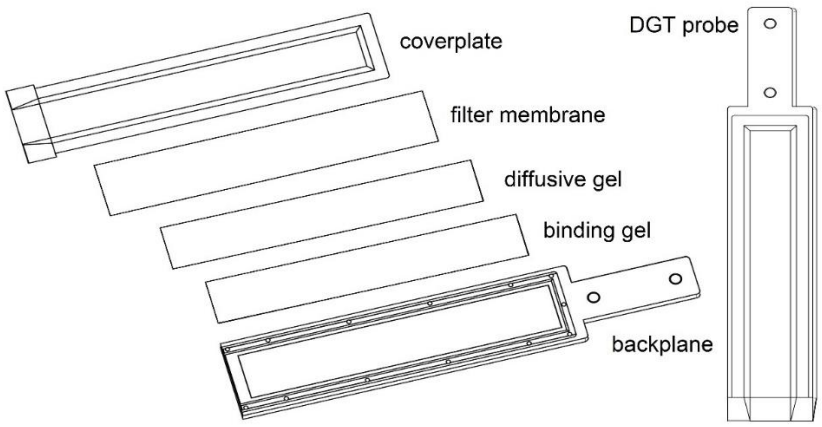


Figure 2.2: The DGT probe structure.

2.5.3 Influencing factors for labile P variability

Studies have indicated that the internal P release is strongly affected by pH, temperature, microbial activity, bioturbation, dissolved reactive P in pore water, resuspension as well as redox condition (Su et al., 2014). Competitive exchange of phosphate anions with OH^- liberates significant amounts of soluble P into water column in high pH condition (Niemistö et al., 2011). High temperature is favorable for the release of P, because elevated temperature strengthens organic mineralization, microbial metabolic processes and solute chemical diffusion (Liu et al., 2018(b); Xue and Lu, 2015). Microbial activity enhances the degradation of organic matter, releasing P from sediments and dead plants (Wang et al., 2013). Worm bioturbation decreased the concentration of labile P in sediment in the first month, then the effects disappeared and emerged again (Chen et al., 2016(b)). The elevated P-adsorbed suspended particulate matter has a negative impact on post-dredging SWI and consequently increases internal P loading (Liu et al., 2016(b)). Huang et al. (2016) have pointed out that the bottom sediment of Lake Taihu could be notably resuspended by wind-driven waves, then enhancing the release of P.

Redox condition is the most popular explanation for controlling P mobilization. It is widely accepted that Fe is the most prevalent redox-sensitive metal and Fe redox cycle plays an important role in P lability. In the oxidized condition, P combines to Fe(III) and settle down to sediments. While in anoxic condition, Fe(III) is reduced to Fe(II) and subsequently both Fe(II) and phosphate are dissolved to water column (Han et al., 2015). Moreover, dissolved sulphate is reduced to S(II) under anoxic circumstances which can combine Fe(II) and form insoluble iron-sulphide (FeS and FeS_2) rapidly (Sun et al., 2016). Therefore, the interactions among P, Fe and S can influence their availability and mobility.

2.6 The GIS application of Al, Ca and Pb in soils

The GIS techniques have been widely used to identify and evaluate metal pollution (Hou et al., 2017). The Pb is one of potential toxic metals which is affected by anthropogenic factors. Otherwise, some metals, such as Al, iron (Fe), are dominated by PMs. In this study, GIS techniques are used to identify geogenic and anthropogenic factors on different spatial distribution patterns of Al, Ca and Pb in urban topsoil of the GLA area.

2.6.1 The different spatial distribution patterns of Al, Ca and Pb

Since Craul (1985) defined urban soils to have "characteristics unlike those of their natural counterparts", the PMs still have major natural control on soil geochemistry even in intensively urbanized areas (Argyraki and Kelepertzis, 2014). Therefore, it is of great interest to investigate the different levels of the natural and anthropogenic influence factors on spatial distributions. Wilcke et al. (1998) found that the metal groups of Al, Chromium (Cr), Fe, manganese (Mn) and nickel (Ni) were dominated by the PMs, whereas the metal groups of cadmium (Cd), copper (Cu), Pb, and zinc (Zn) likely originated from anthropogenic emissions in Bangkok topsoils. Appleton et al. (2013) reported that the variance of Al, Ca and Pb in London region was controlled by soil PMs in dominant, moderate and slight level, respectively. The Al is the most enriched metal in the crust and is existing as oxides or silicates in the mineral feldspars, beryl, and garnet. The Al and Fe are significantly correlated with the clay content (Wilcke et al., 1998). Li et al. (2013(b)) demonstrated that the roadside and residential areas had greater concentrations of Ca, Cu, Mn and Zn compared to other land use types, which is in accordance with Appleton et al. (2013) and Liu et al., (2015). Low background value of Pb in soil due to naturally occurring soil lead minerals such as anglesite, cerussite, and galena (Minca and Basta, 2013). It is believed that the elevated Pb in urban soils is caused by human activity, coming from the historic use of Pb-based products (Wong and Li, 2004).

2.6.2 The Pb contamination in urban soil and its hazard

The Pb contamination of urban soils is pervasive and detrimental to human health, resulting from that soil is a major sink for anthropogenic metals via varied pathways (Deng et al., 2016). Quite a lots cities have been revealed with elevated Pb in soils, for instance, Moscow (Russia) (Nikiforova and Kosheleva, 2007), Sydney (Australia) (Walter et al., 2005), Shenyang (China) (Li et al., 2013(c)), Toronto (Canada) (Nazzal et al., 2015), New Orleans (USA) (Laidlaw et al., 2017). This contamination is mostly derived from Pb-added petroleum, deterioration of Pb-containing paint, emission from Pb smelters and battery recycling and other industries. Although Pb has been banned to be used in petrol in most countries, Pb-based paints, Pb-acid batteries and Pb-soldered printed circuit boards in electronic appliances are still produced in many countries (Sung and Park, 2018). The spatial pattern of contamination is influenced by historical and current land use, traffic, landscape and socioecological landscape (Laidlaw et al., 2017).

Eating edible vegetation is an important pathway to transfer soil Pb to human beings. Urban agriculture has grown in large cities, such as New York City (USA), Oakland (USA), Daejeon (Korea), because it provides many advantages, such as the production of fresh and health foods using less chemical pesticide, and the improvement of physical and mental health of urban residents (Sung and Park, 2018). However, a large portion of vegetables exceeded the local standard for maximum Pb threshold value, presenting a risk of Pb exposure (McBride et al., 2014; McClintock, 2012; Sung and Park, 2018). There were 738 allotments in London and each allotment is usually around 250 m^2 (GLA, 2006). They can be rented by an individual for growing fruit and vegetables. Therefore, the Pb concentration is very crucial for urban farmers.

Children are more amenable than adults to Pb poisoning due to typical childhood hand-to-mouth characteristics (Sahmel et al., 2015). It is recorded in many studies that elevated Pb concentration in children's blood has a harmful impact on neurological and physiological

development (Deng et al., 2016). In the Pearl River Delta region of China, a high number of children have elevated blood Pb level (Chen et al., 2012). The blood Pb levels of preschool children living suburb areas was higher than those who were living in the central district in Shanghai due to the industrial areas moving form city center to suburb areas. This study also pointed out that boys' blood Pb levels were greater than those of girls which possibly caused by more outdoor playing among boys (Cao et al., 2014). Despite the sharp drop of children's blood Pb levels owing to banned Pb-additive gasoline, there is abundant evidence that low-level Pb exposure is associated with intellectual deficits, behavioral impairments, conduct disorder even criminal behavior without apparent threshold (Oulhote et al., 2013). Oulhote et al. (2013) also demonstrated that the steepest increase in blood Pb levels happened at the lowest levels of Pb-contaminated floor dust, which suggested that Pb contamination should be kept as low as possible.

2.7 The GIS application in transferring point data to areal data

2.7.1 Causes for different spatial supports

Spatially misaligned data are becoming more and more ubiquitous because of the development in data collection as well as the increasing scientific interdisciplinary researches, including the geographical, environmental, epidemiological, ecological and economical fields. Data are primarily collected either at locations in space (point data), or over areal units like postal zip zones, residential communities, administrative districts or counties (areal data) (Moraga et al., 2017). For example, concentration of heavy metal in soils, air pollutions, such as sulfur dioxide (SO_2), measured at air monitoring stations, temperature and precipitation estimated at weather stations are often collected as point data, whereas socio-economic data, such as education levels, ethnicity and income, are commonly recorded by census or survey in designed regions. In epidemiology, cases of disease occurrence or death can be recorded with location, but areal data are frequently adopted by aggregating the point data over spatial units due to ethical concerns and patient confidentiality (Lawson, 2012). In practice, we wish to find out the

relationship between two variables which are recorded at different spatial resolution. In the literature, the incompatibility problem is called as point areal spatial misalignment (Finley et al., 2014), though Gotway and Young (2002) prefer the term of "change of support problem".

2.7.2 Methods to address the challenge

The examples of three groups of methodology to address the problem of point areal spatial misalignment are listed (Table 2.1). The first is upscaling approaches which are more often applied (Fig. 2.3). It is the determination of representative values over the area units using the point data. On the opposite direction, downscaling methods is to change the areal data to match the point data (Fig. 2.3). The third one is the jointly model using the two variables, including their spatial variations. There are a variety of approaches, ranging from the straightforward arithmetic average to the more demanding Bayesian models. There is no such a thing as a 'best" approach for all situations. Rather, the user should select an approach according to the specific objective of the study (Goovaerts, 1999).

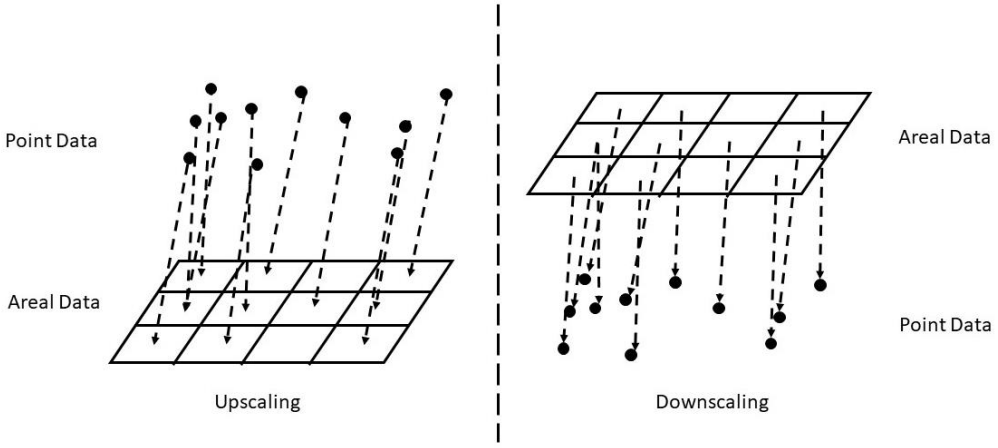


Figure 2.3: Simplified graphical representation of upscaling and downscaling processes between point and areal data.

Table 2.1: Examples of dealing with point areal spatial misalignment.

Group	Methodology	Advantage	Disadvantage	Reference
Upscaling	Arithmetic average/median value Geometric average/median value	easy calculation and application	not suitable with sparse sampling or uneven random sampling	Cave et al., 2018
	Average/median value extracted by interpolation (e.g. IDW, kriging; an empirical model)	unbiased prediction; taking spatial variability into account	the uncertainty and estimated errors with predicted values	Pinichka et al., 2017
	Average/median value extracted by block interpolation (e.g. IDW, kriging)	unbiased prediction; taking spatial variability into account	the uncertainty and estimated errors with predicted values	Rapant et al., 2015
	Centroid value extracted by interpolation (e.g. IDW, kriging; an empirical model)	taking spatial variability into account	biased prediction based on one point	Young et al., 2009
Downscaling	Area-to-point kriging	including covariates and providing standard errors for the resulting prediction	no implantation in GIS software and no defaults that allow automatic solutions	Wang et al., 2016
Modelling	Bayesian models (e.g. Markov Monte Carlo; Integrated Nested Laplace approximation)	taking spatial variation and uncertainty into account	complex model; long computational time	Moraga et al., 2017

2.8 Summary

The literature review indicates: i) It is necessary to investigate the magnitude and location of P accumulation and their sources which is the basement to improve P management and to prevent P from transporting into aquatic systems; ii) The GIS can be used to reveal spatial distribution patterns of P, Al, Ca and Pb in extensively urbanized area; iii) The GIS can be used to describe 2D spatial pattern obtained by the DGT measurement that the classical statistical description cannot satisfy; iv) The different level of geogenic and anthropogenic influences on different spatial distribution patterns of Al, Ca and Pb in urban topsoils of the GLA can be identified using GIS; v) The GIS can be used to transform point data to areal data and different GIS-based methods have respective pros and cons.

Reference

Abbasi, M. K., Musa, N., Manzoor, M., 2015. Mineralization of soluble P fertilizers and insoluble rock phosphate in response to phosphate-solubilizing bacteria and poultry manure and their effect on the growth and P utilization efficiency of chilli (Capsicum annuum L.) Biogeosciences 12, 4607-4619.

Addiscott, T. M., Thomas, D., 2000. Tillage, mineralization and leaching: phosphate. Soil Till. Res. 53, 255-273.

Appleton, J. D., Johnson, C. C., Ander, E. L., Flight, D. M. A., 2013. Geogenic signatures detectable in topsoils of urban and rural domains in the London region, UK, using parent material classified data. Appl. Geochem. 39, 169-180.

Argyraki, A., Kelepertzis, E., 2014. Urban soil geochemistry in Athens, Greece: The importance of local geology in controlling the distribution of potentially harmful trace elements. Sci. Total Environ. 482-483, 366-377.

Barik, S. K., Bramha, S., Bastia, T. K., Behera, D., Mohanty, P. K., Rath, P., 2019. Distribution of geochemical fractions of phosphorus and its ecological risk in sediment cores of a largest brackish water lake, South Asia. Int. J. Sediment Res. 34(3), 251-261.

Bateman, A., Horst, D., Boardman, D., Kansal, A., Carliell-Marquet, C., 2011. Closing the phosphorus loop in England: The spatio-temporal balance of phosphorus capture from manure versus crop demand for fertilizer. Resour. Conserv. Recy. 55(12), 1146-1153.

Bennett, E. M., Carpenter, S. R., Clayton, M. K., 2004. Soil phosphorus variability: scale-dependence in an urbanizing agricultural landscape. Landscape Ecol. 20, 389-400.

Biswas, J. K., Hazra, S., Majumdar, J., Mandal, S. K., Shaheen, S. M., Sarkar, S. K., Meissner, R., Meers, E., Rinklebe, J., 2017. Impact of raking and bioturbation-mediated ecological manipulation on sediment–water phosphorus diagenesis: a mesocosm study supported with radioactive signature. Environ, Geochem. Hlth. 39(6), 1563-1581.

Bundy, L. G., Sturgul, S. J., 2001. A phosphorus budget for Wisconsin cropland. J. Soil Water Conserv. 56, 243-249.

Cao, J., Li, M., Wang, Y., Yu, G., Yan, C., 2014. Environmental lead exposure among preschool children in Shanghai, China: blood lead levels and risk factors. PLoS ONE 9(12), e113297. https://doi.org/10.1371/journal.pone.0113297

Cave, M., Wragg, J., Lister, R., 2018. The effect of lead in soil on crime deprivation in Derby, Leicester and Nottingham. Appl. Geochem. 88, 198-212.

Chan, C., Guisasola, A., Baeza, J. A., 2017. Enhanced Biological Phosphorus Removal at low Sludge Retention Time in view of its integration in A-stage systems. Water Res. 118, 217-226.

Chen, J., Tong, Y., Xu, J., Liu, X., Li, Y., Tan, M., Li, Y., 2012. Environmental lead pollution threatens the children living in the Pearl River Delta region, China. Environ. Sci. Pollut. R. 19(8), 3268-3275.

Chen, M., Ding, S., Chen, X., Sun, Q., Fan, X., Lin, J., Ren, M., Yang, L., Zhang, C., 2018. Mechanisms driving phosphorus release during algal blooms based on hourly changes in iron and phosphorus concentrations in sediments. Water Res. 133, 153-164.

Chen, M., Ding, S., Liu, L., Wang, Y., Xing X., Wang, D., Gong, M., Zhang, C., 2016(b). Fine-scale bioturbation effects of tubificid worm (Limnodrilus hoffmeisteri) on the lability of phosphorus in sediments. Environ. Pollut. 219, 604-611.

Chen, X., Li, H., Hou, J., Cao, X., Song, C., Zhou, Y., 2016 (a). Sediment-water interaction in phosphorus cycling as affected by tropic states in a Chinese shallow lake (Lake Donghu). Hydrobiologia 776(1), 19-33.

Civan, A., Worrall., F., Jarvie, H. P., Howden, J. K., Burt, T. P., 2018. Forty-year trends in the flux and concentration of phosphorus in British rivers. J. Hydrol. 558, 314-327.

Cordell, D., Drangert J., White, S., 2009. The story of phosphorus: global food security and food for thought. Global Environ. Change 19, 292-305.

Costa, M.G., Gama-Rodrigues, A.C., Gonçalves, J.L.M., Gama-Rodrigues, E.F., Sales, M.V.S., Aleixo, S., 2016. Labile and Non-Labile Fractions of Phosphorus and Its

Transformations in Soil under Eucalyptus Plantations, Brazil. Forests 7, 15. https://doi.org/10.3390/f7010015

Craul, P. J., 1985. A description of urban soils and their desired characteristics. J Arboricult 11, 330-339.

Cromley, E. K., McLafferty, S. L., 2002. GIS and public health. The Guildford Press, New York, NY, USA, pp.340.

Davison, W., Zhang, H., 1994. In situ speciation measurements of trace components in natural waters using thin-film gels. Nature 367, 546-548.

Deng, W., Li, X., An, Z., Yang, L., 2016. Lead contamination and source characterization in soils around a lead-zinc smelting plant in a neat-urban environment in Baoji, China. Arch. Environ. Contam. Toxicol. 71, 500-508.

Dewaelheyns, V., Elsen, A., Vandendriessche, H., Gulinck, H., 2013. Garden management and soil fertility in Flemish domestic gardens. Landscape Urban Plan. 116, 25-35.

Ding, S. M., Jia, F., Xu, D., Sun, Q., Zhang, L., Fan, C. X., Zhang, C. S., 2011. High-resolution, two-dimensional measurement of dissolved reactive phosphorus in sediments using the diffusive gradients in thin films technique in combination with a routine procedure. Environ. Sci Technol. 45, 9680-9686.

Finley, A. O., Banerjee, S., Cook, B. D., 2014. Bayesian hierarchical models for spatially misaligned data in R. Methods Ecol. Evol. 5(6), 514-523.

Fu, W., Tunney, H., Zhang, C., 2010. Spatial variation of soil test phosphorus in a long-term grazed experimental grassland field. J. Plan Nutri. Soil Sci. 173, 323-331.

Fu, W., Zhao, K., Zhang, C., Tunney, H., 2011. Using Moran' I and geostatistics to identify spatial patterns of soil nutrients in two different long-term phosphorus-application plots. J. Plan Nutri. Soil Sci. 174, 785-798.

Gao, R., Liu, S., Zhang, Y., Li, H., Huang, Y., Xia, X., Jiang, T., Zhang, H., 2011. Temporal-spatial variability and fractal characteristics of soil nitrogen and phosphorus in Xinji District, Hebei Province, China. Environ. Monit. Assess. 174, 229-240.

García-Ayllón, S., 2017. Diagnosis of complex coastal ecological systems: Environmental GIS analysis of a highly stressed Mediterranean lagoon through spatiotemporal indicators. Ecol. Indic. 83, 451-462.

GLA (Greater London Authority), 2006. A lot to lose: London's disappearing allotments. London: Environment Committee, GLA City Hall.

Goovaerts, P., 1999. Geostatistical tools for deriving block-averaged values of environmental attributes, geographic information sciences. A Journal of the Association of Chinese Professionals in Geographic Information Systems 5(2), 88-96.

Gotway, C. A., Young, L. J., 2002. Combing incompatible spatial data. J. Am. Stat. Assoc. 97(458), 632-648.

Grande, J. D., Karthikeyan, K. G., Miller, P. S., Powell, J. M., 2005. Corn residue level and manure application timing effects on phosphorus losses in runoff. J. Environ. Qual. 34, 1620-1631.

Granger, S. J., Heaton, T. H. E., Pfahler, V., Blackwell, M. S. A., Yuan, H., Collins, A. L., 2017. The oxygen isotopic composition of phosphate in river water and its potential sources in the Upper River Taw catchment, UK. Sci. Total Environ. 574, 680-690.

Guan, F., Xia, M., Tang, X., Fan, S., 2017. Spatial variability of soil nitrogen, phosphorus and potassium contents in Moso bamboo forests in Yong'an City, China. Catena 150, 161-172.

Gupta, N., Aktaruzzaman, M., Wang, C.,2012. GIS-based Assessment and Management of Nitrogen and Phosphorus in Rönneå River Catchment, Sweden. J. Indian Soc. Remote Sens. 40(3), 457-466.

Halliwell, D. J., McKelvie, I. D., Hart, B. T., Dunhill, R. H., 2001. Hydrolysis of triphosphate from detergents in a rural waste water system. Water Res. 35(2), 448-454.

Han, C., Ding, S., Yao, L., Shen, Q., Zhu, C., Wang, Y., Xu, D., 2015. Dynamics of phosphorus–iron–sulfur at the sediment–water interface influenced by algae blooms decomposition. J. Hazard. Mater. 300, 329-337.

He, Q., Rajendran, A., Gan, J., Lin, H., Felt, C. A., Hu, B., 2019. Phosphorus recovery from dairy manure wastewater by fungal biomass treatment. Water Environ. J. 33, 508-517.

Herzel, H., Krüger, O., Hermann, L., Adam, C., 2016. Sewage sludge ash – a promising secondary phosphorus source for fertilizer production. Sci. Total Environ. 542B, 1136-1143.

Hou, D., O'Connor D., Nathanail, P., Tian, L., Ma, Y., 2017. Integrated GIS and multivariate statistical analysis for regional scale assessment of heavy metal soil contamination: A critical review. Environ. Pollut. 231, 1188-1200.

Hu, X., Chen, F., Nagle, G., Fang, Y., Yu, M., 2011. Soil phosphorus fractions and tree phosphorus resorption in pine forests along an urban-to-rural gradient in Nanchang, China. Plant Soil 346, 97-106.

Hu, Z., Haneklaus, S., Liu, Q., Xu, C., Cao, Z., Schnug, E., 2003. Small-scale spatial variability of phosphorus in a paddy soil. Commun. Soil Sci. Plan. 34(19-20), 2791-2801.

Huang, L., Fang, H., He, G., Jiang, H., Wang, C., 2016. Effects of internal loading on phosphorus distribution in the Taihu Lake driven by wind waves and lake currents. Environ. Pollut. 219, 760-773.

Iho, A., Laukkanen, M., 2012. Precision phosphorus management and agricultural phosphorus loading. Ecol. Econ. 77, 91-102.

Jia, W., Yan, Z., Chadwick, D., Kang, L., Duan, Z., Bai, Z., Chen, Q., 2015. Integrating soil testing phosphorus into environmentally based manure management in peri-urban regions: a case study in the Beijing area. Agr. Ecosyst. Environ. 209, 47-59.

Kelessidis, A., Stasinakis, A. S., 2012. Comparative study of the methods used for treatment and final disposal of sewage sludge in European countries. Waste Manag. 32(6),1186-1195.

Laidlaw, M. A. S., Filippelli, G. M., Brown, S., Paz-Ferreiro, J., Reichman, S. M., Netherway, P., Truskewycz, A., Ball, A. S., Mielke, H. W., 2107. Case studies and evidence-based approaches to addressing urban soil lead contamination. Appl. Geochem. 83, 14-30.

Lawson, A. B., 2012. Bayesian point event modeling in spatial and environmental epidemiology. Stat. Methods Med. Res. 21, 509-529.

Li, G. L., Bai, X., Sheng, Y., Zhang, H., Zhu, Y. G., 2011. Urban phosphorus metabolism through food consumption: the case of China. J. Ind. Ecol. 16(4), 588-599.

Li, G., Ittersum, M. K., Leffelaar, P. A., Sattari, S. Z., Li, H., Huang, G., Zhang, F., 2016. A multi-level analysis of China's phosphorus flows to identify options for improved management in agriculture. Agr. Syst. 144, 87-100.

Li, H., Liu, J., Li, G., Shen, J., Bergström, L., Zhang, F., 2015. Past, present, and future use of phosphorus in Chinese agriculture and its influence on phosphorus losses. Ambio 44(Suppl. 2), 274-285.

Li, H., Liu, L., Li, M., Zhang, X., 2013(a). Effects of pH, temperature, dissolved oxygen, and flow rate on phosphorus release processed at the sediment and water interface in storm sewer. J. Anal. Methods Chem. vol. 2013, Article ID 104316, 7 pages. https://doi.org/10.1155/2013/104316

Li, S., Yuan, Z., Bi, J., Wu, H., 2010. Anthropogenic phosphorus flow analysis of Hefei city, China. Sci. Total Environ. 408(23), 5715-5722.

Li, X., Liu, L., Wang, Y., Luo, G., Chen, X., Yang, X., Hall, M. H. P., Guo, R., Wang, H., Cui, J., He, X., 2013(c). Heavy metal contamination of urban soil in an old industrial city (Shenyang) in Northeast China. Geoderma 192, 50-58.

Li, Z., Zhang, G., Liu, Y., Wan, K., Zhang, R., Chen, F., 2013(b). Soil Nutrient Assessment for Urban Ecosystems in Hubei, China. PLoS ONE 8(9), e75856. https://doi.org/10.1371/journal.pone.0075856

Liu, C., Shao, S., Shen, Q., Fan, C., Zhang, L., Zhou, Q., 2016(b). Effects of riverine suspended particulate matter on the post-dredging increase in internal phosphorus loading across the sediment-water interface. Environ. Pollut. 211, 165-172.

Liu, C., Zhong, J., Wang, J., Zhang, L., Fan, C., 2016(a). Fifteen-year study of environmental dredging effect on variation of nitrogen and phosphorus exchange across the sediment-water interface of an urban lake. Environ. Pollut. 219, 639-648.

Liu, H., Bu, H., Liu, G., Wang, Z., Liu, W., 2015. Effects of surrounding land use on metal accumulation in environments and submerged plants in subtropical ponds. Environ. Sci. Pollut. Res. 22, 18750-18758.

Liu, Q., Ding, S., Chen, X., Sun, Q., Chen, M., Zhang, C., 2018 (b). Effects of temperature on phosphorus mobilization in sediments in microcosm experiment and in the field. Appl. Geochem. 88, 158-166.

Liu, X., Beusen, A. H. W., Beek, L. P. H., Mogollón, J. M., Ran, X., Bouwman, A. F., 2018(a). Exploring spatiotemporal changes of the Yangtze River (Changjiang) nitrogen and phosphorus sources, retention and export to the East China Sea and Yellow Sea. Water Res. 142, 246-255.

Mage, S. M., Porder, S., 2013. Parent material and topography determine soil phosphorus statues in the Luquillo Mountains of Puerto Rico. Ecosystems 16, 284-294.

McBride, M. B., Shayler, H. A., Spliethoff, H. M., Mitchell, R. G., Marquez-Bravo, L. G., Ferenz, G. S., Russell-Anelli, J. M., Casey, L., Bachman, S., 2014. Concentrations of lead, cadmium and barium in urban garden-grown vegetables: the impact of soil variables. Environ. Pollut. 194, 254-261.

McClintock, N., 2012. Assessing soil lead contamination at multiple scales in Oakland, California: Implications for urban agriculture and environmental justice. Appl. Geogr. 35(1-2), 460-473.

Mcginley, P. M., Masarik, K. C., Gotkowitz, M. B., Mechenich, D. J., 2016. Impact of anthropogenic geochemical change and aquifer geology on groundwater phosphorus concentrations. Appl. Geochem. 72, 1-9.

Meng, Y., Ding, S., Gong, M., Chen, M., Wang, Y., Fan, X., Shi, L., Zhang, C., 2018(a). Submillimeter-scale heterogeneity of labile phosphorus in sediments characterized by diffusive gradients in thin films and spatial analysis. Chemosphere 194, 614-621.

Meng, Y., Cave, M., Zhang, C., 2018(b). Spatial distribution patterns of phosphorus in topsoils of Greater London Authority area and their natural and anthropogenic factors. Applied Geochemistry 88, 213-220.

Metson, G. S., Lin, J., Harrison, J. A., Compton, J. E., 2017. Linking terrestrial phosphorus inputs to riverine export across the United States. Water Res. 124, 177-191.

Millier, H. K. G. R., Hooda, P. S., 2011. Phosphorus species and fractionation - why sewage derived phosphorus is a problem. J. Environ. Manage. 92(4), 1210-1214.

Minca, K. K., Basta, N. T., 2013. Comparison of plant nutrient and environmental soil tests to predict Pb in urban soils. Sci. Total Environ. 445-446, 57-63.

Moore, P. A., Reddy, K. R., Fisher, M. M., 1998. Phosphorus flux between sediment and overlying water in Lake Okeechobee, Florida: spatial and temporal variations. J. Environ. Qual. 27(6), 1428-1439.

Moraga, P., Cramb, S. M., Mengersen, K. L., Pagano, M., 2017. A geostatistical model for combined analysis of point-level and area-level using INLA and SPDE. Spat. Stat. 21, 27-41.

Nazzal, Y. N., Al-Arifi, N. S. N., Jafri, M. K., Kishawy, H. A., Ghrefat, H. A., El-Waheidi, M. M., Batayneh, A. A., Zumlot, T. A., 2015. Multivariate statistical analysis of urban soil contamination by heavy metals at selected industrial location in the Greater Toronto area, Canada. Geologia Croatica 68(2), 147-159.

Niemietz, K., Wagner, A., Gründig-Wendrock, B., Stoyan, D., Niklas, J. R., 2010. Variogram analysis of charge-carrier effective lifetime topograms in mc-Si materials. Sol. Energy Mater. Sol. Cells 94, 164-170.

Niemistö, J., Holmroos, H., Horppila, J., 2011. Water and sediment resuspension regulating internal phosphorus loading in a shallow lake – field experiment on diurnal variation. J. Limnol. 70(1), 3-10.

Nikiforova, E. M., Kosheleva, N. E., 2007.Dynamics of contamination of urban soils with lead in the eastern district of Moscow. Eurasian Soil Sci. 40(8), 880-892.

Oliver, M. A., Webster, R., 2015. Basic steps in geostatistics: the variogram and kriging. ISSN: 2211-808X. Springer, p. 1. https://doi.org/10.1007/978-3-319-15865-5

Oulhote, Y., Tertre, A. L., Etchevers, A., Bot, B. L., Lucas, J., Mandin, C., Strat, Y., Lanphear, B., Glorennec, P., 2013. Implications of different residential lead standards on children's blood lead levels in France: predictions based on a national cross-sectional survey. Int. J. Hyg. Envir. Heal. 216(6), 743-750.

Ouyang, Y., Leininger, T. D., Moran, M., 2015. Estimating effects of reforestation on nitrogen and phosphorus load reductions in the Lower Yazoo River Watershed, Mississippi. Ecol. Eng. 75, 449-456.

Pinichka, C., Makka, N., Sukkumnoed, D., Chariyalertsak, S., Inchai, P., Bundhamcharoen, K., 2017. Burden of disease attributed to ambient air pollution in Thailand: A GIS-based approach. PLoS ONE 12(12), e0189909. https://doi.org/10.1371/journal.pone.0189909

Pretty, J. N., Mason, C. F., Nedwell, D. B., Hine, R. E., 2002. A preliminary assessment of the environmental costs of the eutrophication of fresh water in England and Wales. Commissioned by the Environment Agency; Project Officers: S. Leaf and R. Dils. University of Essex, Colchester, UK.

Pu, P. M., Wang, G. X., Hu, C. H., Hu, W. P., Fan, C. X., 2000. Can we control lake eutrophication by dredging? J. Lake Sci. 12(3), 269-279 (in Chinese).

Rapant, S., FajčÍková, K., Cvečková, V., Ďurža, A., Stehlíková, B., Sedláková, D., Ženišová, Z., 2015. Chemical composition of groundwater and relative mortality for cardiovascular diseases in the Slovak Republic. Environ. Geochem. Health 37, 745-756.

Reynolds C. S., Davies, P. S., 2001. Sources and bioavailability of phosphorus fraction in freshwaters: a British perspective. Biol. Rev. 76, 27-64.

Roger, A., Libohova, Z., Rossier, N., Joost, S., Maltas, A., Frossard, E., Sinaj, S., 2014. Spatial variability of soil phosphorus in the Fribourg canton, Switzerland. Geoderma 217-218, 26-36.

Sahmel, J., Hsu, E. I., Avens, H. J., Beckett, E. M., Devlin, K. D., 2015. Estimation of hand-to-mouth transfer efficiency of lead. Ann. Occup. Hyg. 59(2), 210-220.

Sarvajayakesavalu, S., Lu, Y., Withers, P. J. A., Pavinato, P. S., Pan, G., Chareonsudjai, P., 2018. Phosphorus recovery: a need for an integrated approach. Ecosystem Health and Sustainability, 4(2), 48-57.

Singh, O., Kasana, A., 2017. GIS-based spatial and temporal investigation of groundwater level fluctuation under rice-wheat ecosystem over Haryana. J. Geological Society of India 89, 554-562.

Spears, B. M., Carvalho, L., Perkins, R., Kirika, A., Paterson, D. M., 2006. Spatial and historical variation in sediment phosphorus fractions and mobility in large shallow lake. Water Res. 40(2), 383-391.

Smil, V., 2000. Phosphorus in the environment: natural flows and human interferences. Annu. Rev. Energy Environ. 25, 53-88.

Smith, C., Dawson, D., Archer, J., Davies, M., Frith, M., Hughes, E., Massini, P., 2011. From green to fray; observed changes in garden vegetation structure in London, 1998-2008. London Wildlife Trust, Greenspace Information for Great London, and Great London Authority.

Su, J., Bochove, E. Auclair, J., Thériault, G., Hu, C., Li, X., 2014. Phosphorus fluxes at the sediment-water interface in a temperate region agriculture catchment. Water Air Soil Pollut. 225, 1739. https://doi.org/10.1007/s11270-013-1739-2

Sun, Q., Sheng, Y., Yang, J., Bonito, M. D., Mortimer, R. J. G., 2016. Dynamic characteristics of sulfur, iron and phosphorus in coastal polluted sediments, north China. Environ. Pollut. 219, 588-595.

Sung, C. Y., Park, C., 2018. The effect of site- and landscape-scale factors on lead contamination of leafy vegetables grown in urban gardens. Landscape Urban Plan. 177, 38-46.

Szögi, A. A., Vanotti, M. B., Hunt, P. G., 2015. Phosphorus recovery from pig manure solids prior to land application. J. Environ. Manage. 157, 1-7.

Tang, X., Wu, M., Li, R., 2018. Phosphorus distribution and bioavailability dynamics in the mainstream water and surface sediment of the Three Gorges Reservoir between 2003 and 2010. Water Res. 145, 321-331.

Ulén, B., Bechmann, M., Flster, J., Jarvie, H. P., Tunney, H., 2007. Agriculture as a phosphorus source for eutrophication in the north-west European countries, Norway, Sweden, United Kingdon and Ireland: a review. Soil Use Manage. 23(Suppl. 1), 5-15.

Vemulapalli, S. S., Ulak, M.B., Ozguven, E. E., Sando, T., Horner, M. W., Abdelrazig, Y., Moses, R., 2017. GIS-based spatial and temporal analysis of aging-involved accidents: a case study of three counties in Florida. Appl. Spatial Analysis 10, 537-563.

Walter, C., McBratney, A. B., Rossel, R. A. V., Markus, J. A., 2005. Spatial point-process statistics: concepts and application to the analysis of lead contamination in urban soil. Environmetrics 16, 339-355.

Wang, C., Geng, Y., Cheng, L., Mao, Y., 2018. Speciation, mass loading, and fate of phosphorus in the sewage sludge of China. Environ. Sci. Pollut. Res. 25, 35531-35537.

Wang, H., Holden, J., Spera, K., Xu, X., Wang, Z., Luan, J., Xu, X., Zhang, Z., 2013. Phosphorus fluxes at the sediment-water interface in subtropical wetlands subjected to experimental warming: a microcosm study. Chemosphere 90, 1794-1804.

Wang, J., Wang, Y., Zeng, H., 2016. A geostatistical approach to the change-of-support problem and variable-support data fusion in spatial analysis. J. Geogr. Syst. 18, 45-66.

Wang, Q., Li, Y., Ouyang, Y., 2011. Phosphorus fractionation and distribution in sediments from wetlands and canals of a water conservation area in the Florida Everglades. Water Resour. Res. 47, W05550. https://doi.org/10.1029/2009WR008934

Weaver, D. M., Wong, M. T. F., 2011. Scope to improve phosphorus (P) management and balance efficiency of crop and pasture soils with contrasting P status and buffering indices. Plant Soil 349, 37-54.

Wei, Z., Yan, X., Lu, Z., Wu, J., 2019. Phosphorus sorption characteristics and related properties in urban soils in southeast China. Catena 175, 349-355.

Wilcke, W., Müller, S., Kanchanakool, N., Zech, W., 1998. Urban soil contamination in Bangkok: heavy metal and aluminium partitioning in topsoils. Geoderma 86, 211-228.

Wilson, H. F., Satchithanantham, S., Moulin, A. P., Glenn, A. J., 2016. Soil phosphorus spatial variability due to landform, tillage, and input management: a case study of small watersheds in southwestern Manitoba. Geoderma 280, 14-21.

Wollmann, I., Gauro, A., Müller, T., Möller, K., 2018. Phosphorus bioavailability of sewage sledge-based recycled fertilizers. J. Plant Nutr. Soil Sci. 181, 158-166.

Wong, C. S. C., Li, X. D., 2004. Pb contamination and isotopic composition of urban soils in Hong Kong. Sci. Total Environ. 319(1-3), 185-195.

Woo, H., Acuna, M., Moroni, M., Taskhiri, M. S., Turner, P., 2018. Optimizing the Location of Biomass Energy Facilities by Integrating Multi-Criteria Analysis (MCA) and Geographical Information Systems (GIS). Forests 9(10), 585. https://doi.org/10.3390/f9100585

Wood, T. E., Matthews, D., Vandecar, K., Lawrence, D., 2016. Short-term variability in labile soil phosphorus is positively related to soil moisture in a humid tropical forest in Puerto Rico. Biogeochemistry 127, 35-43.

Wu, J., Franzén, D., Malnström, M. E., 2016. Anthropogenic phosphorus flows under different scenarios for the city of Stockholm, Sweden. Sci Total Environ. 542, 1094-1105.

Wu, Z., Liu, Y., Liang, Z., Wu, S., Guo, H., 2017. Internal cycling, not external loading, decides the nutrient limitation in eutrophic lake: a dynamic model with temporal Bayesian hierarchical inference. Water Res., 116, 231-240.

Xing, X., Ding, S., Liu, L., Chen, M., Yan, W., Zhao, L., Zhang, C., 2018. Direct evidence for the enhanced acquisition of phosphorus in the rhizosphere of aquatic plants: a case study on Vallisneria natans. Sci Total Environ. 616-617, 386-396.

Xu, D., Chen, Y., Ding, S., Sun, Q., Wang, Y., Zhang, C., 2013. Diffusive gradients in thin films technique equipped with a mixed binding gel for simultaneous measurements of dissolved reactive phosphorus and dissolved iron. Environ. Sci. Technol. 47, 10477-10484.

Xu, D., Wu, W., Ding, S., Sun, Q., Zhang, C., 2012. A high-resolution dialysis technique for rapid determination of dissolved reactive phosphate and ferrous iron in pore water of sediments. Sci. Total Environ. 421-422, 245-252.

Xu, G., Li, Z., Li, P., Zhang, T., Cheng, S., 2014. Spatial variability of soil available phosphorus in a typical watershed in the source area of the middle Dan River, China. Environ. Earth Sci. 71, 3953-3962.

Xu, H., Demetriades, A., Reimann, C., Jiménez, J. J., Filser, J., Zhang, C., GEMAS Project Team, 2019. Identification of the co-existence of low total organic carbon contents and low pH values in agricultural soil in north-central Europe using hot spot analysis based on GEMAS project data. Sci. Total Environ. 678, 94-104.

Xue, W., Lu, S., 2015. Effects of inactivation agents and temperature on phosphorus release from sediment in Dianchi Lake, China. Environ. Earth Sci. 74(5), 3857-3865.

Xue, Y., Wen, Z., Ji, X., Bressers, H. T. A., Zhang, C., 2017. Location optimization of urban mining facilities with maximal covering model in GIS: a case of China. J. Ind. Ecol. 21(4), 913-923.

Yao, Y., Wang, P., Wang, C., Hou, J., Miao, L., Yuan, Y., Wang, T., Liu, C., 2016. Assessment of mobilization of labile phosphorus and iron across sediment-water interface in a

shallow lake (Hongze) based on in situ high-resolution measurement. Environ. Pollut. 219, 873-882.

Yost, R. S., Uehara, G., Fox, R. L., 1982. Geostatistical analysis of soil chemical properties of large land areas: I. Semi-variograms. Soil Sci. Soc. Am. J. 46, 1028-1032.

Young, L. J., Gotway, C.A., Yang, J., Kearney, G., DuClos, C., 2009. Linking health and environmental data in geographical analysis: It's so much more than centroids. Spatial Spatio-temporal Epidemiol. 1, 73-84.

Yu, J., Ding, S., Zhong, J., Fan, C., Chen, Q., Yin, H., Zhang, L., Zhang, Y., 2017. Evaluation of simulated dredging to control internal phosphorus release from sediments: focused on phosphorus transfer and resupply across the sediment-water interface. Sci. Total Environ. 592, 662-673.

Yuan, Z., Ji, J., Sheng, H., Jiang, S., Chen, T., Liu, X., Liu, X., Zhuang, Y., Zhang, L., 2019. Animal based diets and environment: perspective from phosphorus flow quantifications of livestock and poultry raising in China. J. Environ. Manage. 244(15), 199-207.

Zhang, C., Ding, S., Xu, D., Tang, Y., Wong, M. H., 2014. Bioavailability assessment of phosphorus and metal in soils and sediments: a review of diffusive gradients in thin films (DGT). Environ. Monit. Assess. 186, 7367-7378.

Zhang, G., Burghardt, W., Lu, Y., Gong, Z-T., 2001. Phosphorus-enriched soils of urban and suburban Nanjing and their effect on the groundwater phosphorus. J. Plant Nutr. Soil Sci. 164, 295-301.

Zhang, H., Tripathi, N. K., 2018. Geospatial hot spot analysis of lung cancer patients correlated to fine particulate matter ($PM_{2.5}$) and industrial wind in Eastern Thailand. J. Clean. Prod. 170, 407-424.

Zhang, M. K., 2004. Phosphorus accumulation in soils along an urban-rural land use gradient in Hangzhou, Southeast China. Commun. Soil Sci. Plant Anal. 35, 819-833.

Zhang, Z., Yu, X., Qian, S., 2010. Spatial variability of soil nitrogen and phosphorus of a mixed forest ecosystem in Beijing, China. Environ. Earth Sci. 60, 1783-1792.

Chapter 3
Material and Methodology

3.1 Study area

3.1.1 The GLA area, the UK

The GLA area had a 1579 km² administrative area and a population of more than 8.67 million in 2015 (Lu et al., 2018). London, one of the largest urban zones, has a two-thousand-year history as a major settlement. The properties of urban soils could be changed considerably under substantial human activities. Recently, urban soil is receiving more and more attentions due to its close relationship with environment and potential association with public health (Li et al., 2018).

A simplified geological map (Miles and Appleton, 2005) (Fig. 3.1) presents a bedrock sequence from Cretaceous Period to Paleogene Period of the GLA area. The bedrock sequence of Cretaceous Period ranges from Lower Greensand Group, Selborne Group (Gault Formation and Upper Greensand Formation), Grey Chalk Group to the top of White Chalk Group. Chalk crops out in the south part of the GLA area which is typically a fine-grained white limestone and contains almost pure calcium carbonate. The Paleogene Period bedrock includes Lambeth group, Thames group and Bracklesham group. In addition, some parts of the area are covered by considerable Quaternary superficial deposits (BGS, 2011). The Quaternary deposits consist of clay-with-flints, plateau gravels, glacial till, river terrace deposits, brickearth, alluvium, and Head. The clay-with-flints are derived from the original Palaeogene cover via weathering and solifluction. Glacial till is a deposition of the Anglian ice sheet which is primarily made of pebbles or boulder-rich clay. River terrace deposits took the current shape due to the diversion of the River Thames. The main PMs of flat surface in valleys of the rivers Thames and Lee are alluvial deposits, and they contain principally silty clay and clayey-silt with subsidiary sands. Brickearth is loessic deposits with basal gravel.

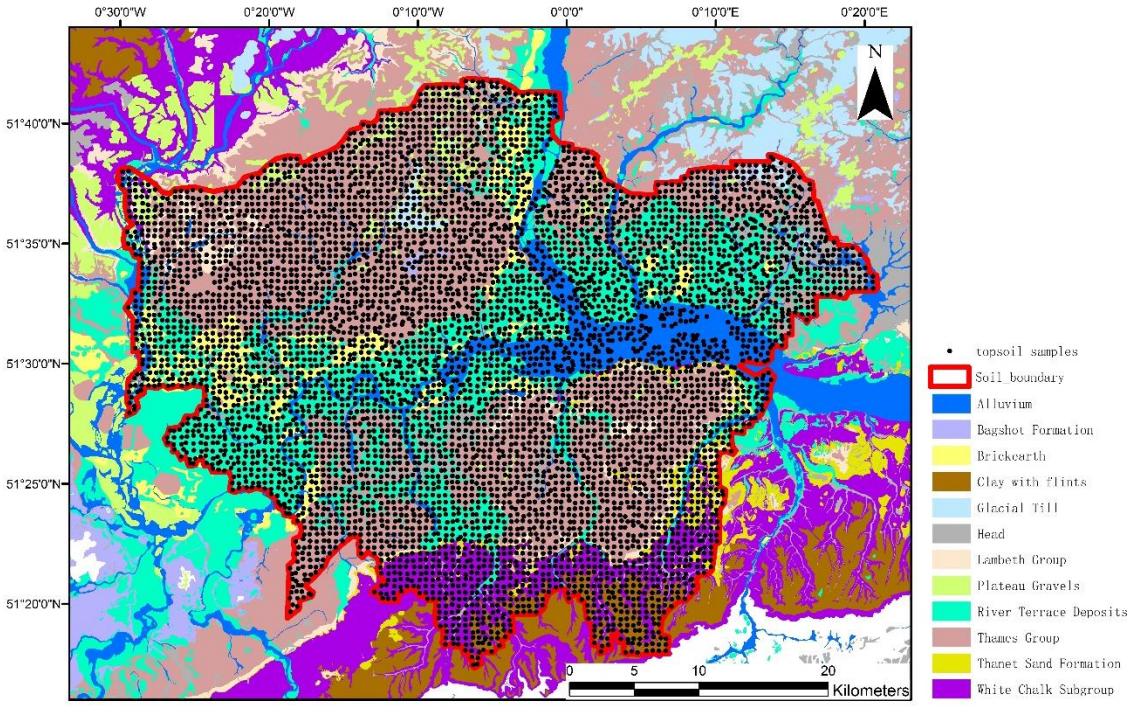


Figure 3.1: Simplified geology map of the GLA area overlaid with soil sampling locations.

3.1.2 Lake Taihu, China

Lake Taihu (30°25'40" - 31°32'58" N, 119°52'32" - 120°36'10" E) is the third biggest freshwater lake of China. It lies on the Yangtze River delta which is one of the most developed region in China. It has a total annual average surface area of 2338 km^2 and average depth of 1.89 m (Gao et al., 2006). The whole lake consists of Zhushan Bay, Meiliang Bay, Gonghu Bay, Eastern Bays and central area (Fig. 3.2). Over the past few decades, large quantities of nutrients' inputs have made it eutrophic. Several times of severe cyanobacterial bloom have happened in Lake Taihu in the recent years, especially in west and north regions of Lake Taihu. The disastrous consequence of cyanobacterial bloom in 2007 was the shortage of drinking water for more than 2 million people, because Lake Taihu is one of the primary water sources for a few large cities including Suzhou, Wuxi and Huzhou (Wang et al., 2016).

Material and Methodology

The sampling site is situated in Meiliang Bay, northwest region of Lake Taihu, which has experienced several times of cyanobacterial blooms. Meiliang Bay receives a huge amount of residential and industrial wastewater from three major inflowing rivers, the Wujin Gang, Zhihu Gang and Liangxi River (Ye et al., 2015). Although industrial and domestic sewage have been largely controlled, internal P loading can still persist for 5-15 years (Ding et al., 2018).

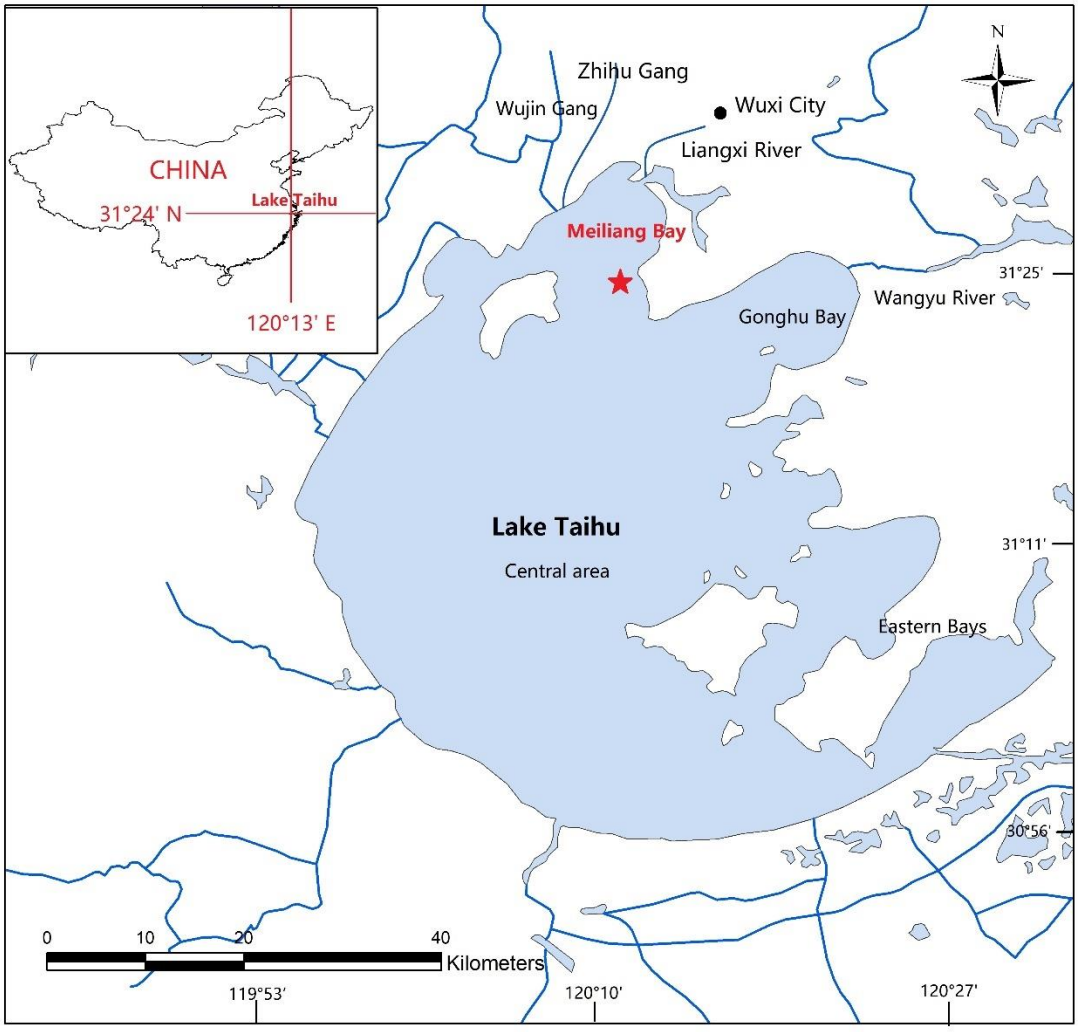


Figure 3.2: Sketch map of Lake Taihu.

3.2 Sampling and laboratory analysis

3.2.1 Soil sampling at the GLA area

A total of 6467 topsoil samples were taken using a hand-held auger by British Geological Survey (BGS) during 2005 and 2009 in the GLA area. The sampling was collected with a depth of 5-20 cm and each sample was gained on the node of a 1 km × 1 km grid (Fig. 3.1). In the purpose of reducing the influence of small-scale heterogeneity of the chemical elements, each sample comprised 5 sub-samples, taken from the center and four corners of a 20 m × 20 m square. After drying, the soil samples were sieved through a 2 mm nylon mesh, with the aim to remove stones and plant debris. Then, the samples were simply ground by hand in a pestle and mortar. Next, the samples were sealed into vessels which contained mill balls. Finally, after adding solid binder to improve adhesion of soil samples, the milled powder was compressed into pellets (Johnson, 2011). Forty-eight chemical elements, including Al, Ca, Pb and P, were examined by X-ray fluorescence (XRF) spectrometry. Moreover, loss on ignition (LOI at 450 ºC) and pH were determined. More details about sample preparation, analytical methods, and quality control procedures were described in Allen et al. (2011) and Johnson (2011).

3.2.2 Sediment sampling at Lake Taihu

Spanning from February 2016 to January 2017, colleagues in Nanjing collected three parallel sediment cores into plexiglass tubes using a gravity core sampler monthly at the sampling site of Meiliang Bay (Fig. 3.2). They transported the core sediments to laboratory and deposited them overnight, while the temperature was maintained at the same temperature as in the field. Then, one Zirconium-oxide DGT probe (Fig. 2.2) which had been already deoxygenated with N for 16 hours was put into each core vertically and deployed for an entire day (24 hours). The DGT probes applied in this study were made of high-capacity Zirconium-oxide binding gel, which were produced in the guide of the procedure described by Ding et al. (2011). A

polyvinylidene difluoride (PVDF) filter membrane with a pore size of 0.45 μm and a thickness of 0.10 mm bought from Millipore Corporation was chosen as the diffusion layer for the purpose of submillimeter measurement. After retrieval, the SWI was drawn immediately. Subsequently, each probe was rinsed using deionized water. The accumulated mass of P in the binding gel was measured referring to the method reported by Ding et al. (2013). Each binding gel was heated in hot water of 85 ℃ for 5 days. Afterwards, it was immersed in a vessel with the mixed reagent (ammonium molybdate tetrahydrate and potassium antimonyl tartrate), kept at 35 ℃ for 45 minutes. In the end, the surface of the binding gel was scanned by a flat-bed scanner at a resolution of 600 dpi, corresponding to a pixel size of 42 μm × 42 μm. The grayscale intensity was calculated into mass of P accumulated per area (μg cm^{-2}) according to Equation (3.1) (Ding et al., 2013):

$$G = -167.29e^{\frac{M}{-6.51}} + 214.63 \quad (3.1)$$

where G is the grayscale intensity of the gel surface, and M (μg cm^{-2}) is the accumulated mass over the deployment time (24 hours). In this case, the measured concentration of labile P in sediments was far lower than the concentration in the pore water. Therefore, the DGT-labile P was represented as flux (pg cm^{-2} s^{-1}) (Widerlund et al., 2012) and the flux was calculated based on Equation (3.2).

$$F = \frac{10^6}{24\times60\times60}M \quad (3.2)$$

3.3 Data analysis

3.3.1 Descriptive statistics

3.3.1.1 Representatively descriptive parameters

Before any treatment or analysis on a data set, we carried out the descriptive statistics to obtain its basic information. The descriptive parameters include sampling number, average,

percentiles, standard deviation, coefficient of variation, skewness, etc. The arithmetic mean (average), median and mode are three measures of centrality. The mean summaries all the information of the observations and is the center of mass of the data set. The median is an observation that lies in the middle of a data set. Compared to the mean, the median is less affected by extreme observations. The variance and standard deviation are frequently used to measure variability or dispersion. The larger the variance and standard deviation means that the farther deviation of the observations from their mean (Aczel, 1995). The coefficient of variance is a normalized measure of dispersion which is useful for comparing data sets with different units or means (Gallardo and Paramá, 2007). The skewness is a descriptor of the asymmetry of the probability distribution of a data set. Under common condition, negative skewness suggests left-tailed, while positive skewness suggests right-tailed (Groeneveld and Meeden, 1984).

3.3.1.2 Frequency distribution graphs

There are plenty of sorts of graphs used to display data directly. We can choose different graphs out of different interests. A histogram commonly exhibits a frequency distribution with aggregation of data into groups. Comparing a histogram to a normal probability curve is an informal but straightforward method for testing normality, especially when the sample is big (Lin and Mudholkar, 1980). A box plot is a comprehensive approach to illustrate the central tendency (the median and mean), spread (the interquartile range), skewness (the position of median) and the outliers (the values beyond the outer fences) (Aczel, 1995). We can draw several box plots to compare several data sets.

3.3.2 Data treatment

3.3.2.1 Data transformation for London data

The normal distribution is crucial for conventional statistics and geostatistics (Clark and Harper, 2000). Unfortunately, soil geochemistry data are usually abnormally distributed. High skewness and outliers of raw data make the mean no longer the best representative of central tendency, which would make the results of ANOVA unreliable (Aczel, 1995). Additionally, abnormal distributed data sets would also lead to unstable estimates in geostatistics (Webster and Oliver, 2007). For example, the extreme value would affect the calculation of Moran's I and local Moran's I through the influence of the mean value. Therefore, data transformation is essential to normalize the raw data set with abnormal distribution. Logarithmic transformation is one of the most popular ways and is often used to transfer positively skewed data sets to normal distribution (Webster and Oliver, 2007; Zhang et al., 2008). In our study, the logarithmic transformation was applied for dealing with the data sets which needed to be transformed.

3.3.2.2 Data analyses for Lake Taihu data

The majority of the grayscale images scanned from binding gels of DGT comprised 472 × 2829 pixels with the size of 42 μm × 42 μm. There were slightly differences in the number of pixels among the biding gels owning to some errors happened during the experiment procedure. To get the values of greyscale intensity, the raster data of grayscale images were changed to point data using ArcGIS (version 10.3). A point feature class of 1,335,288 (472 × 2829) greyscale intensity values was obtained from each image. The labile P flux was calculated from the greyscale intensity. The basic descriptive statistics of each image were computed using the whole original values of P flux, while the semivariograms and Moran's I analysis were conducted based on the aggregated values of P flux. It is because that the computer cannot deal with the calculation of the original point data which were too large. The aggregated data were generated using "raster" package in free R program (version 3.4.1). The aggregation of 10×10 points into 1 point was based on the mean value. The new size of each aggregated data set was 13584 (48 × 283) with a resolution of 0.42 mm × 0.42 mm. The semivariograms were obtained in ArcGIS. The Moran's I was computed using spatial autocorrelation tool

under spatial statistics tools in ArcGIS. The conceptualization of spatial relationship of this tool was chosen as "Inverse Distance" which means that the nearby neighboring values carry larger weight than the values that are further away. Selecting a distance band value as 10 mm (the minimum of the range) in this case. The spatial correlograms were drawn by free R program. The spatial distribution maps of P flux were produced using ArcGIS software.

3.3.3 One-way ANOVA

The ANOVA was applied to detect whether there were significant differences in the chosen chemical elements among PMs or land use groups in the London studies. The ANOVA statistics are based on two assumptions: (i) the population follows normal distribution; (ii) the observations are random and independent (Aczel, 1995). But these two requirements are seldom satisfied entirely by geochemical data because of the demand of regular sample density. Logarithmic transformation (based on 10) was performed to improve the normality to meet the statistical assumption when the skewness coefficient exceeded 1.0 (Appleton et al., 2013). In the specific application of ANOVA, the null hypothesis would be that different PMs or land use groups had the same effect on the variables. When the probability (p-value) was less than the significance level ($p < 0.001$), we rejected the null hypothesis, meaning that the differences observed among PMs or land use groups were unlikely to be caused by random chance.

3.3.4 Spatial analysis

3.3.4.1 Geostatistics

Geostatistics was developed in the 1960s as a methodology for ore reserve evaluation in mining deposits. Until the late 1980s, it became an approach to analyze spatial structure and

interpolate the value of the variable at unsampled places. Recently, it is increasingly applied to model spatial uncertainty which is useful in risk analysis and decision making (Goovaerts, 1997).

(a) Semivariogram

Semivariogram is fundamental to geostatistics, and semivariance can be calculated according to Equation (3.3) (Cressie, 1993):

$$r(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (3.3)$$

where $r(h)$ is the semivariance for a lag distance h between a pair of observations $Z(x_i)$ and $Z(x_i+h)$, and $N(h)$ is the number of pairs of data separated by lag distance h. The semivariogram is a graph of an ordered set of semivariances obtained by changing h.

Theoretically, $r(h)$ is continuous and the semivariogram increases from 0 at h=0. However, in practice, the experimental or sample semivariogram is a graph of fluctuating plotted points, and $r(h)$ is a positive value when h approaches zero (Fig. 3.3). This discrepancy near the origin is called nugget, which consists of measurement error and spatial variation occurred at distances less than the shortest sampling interval (Niemietz et al., 2010). The observed fluctuation includes error and structural increasing. The semivariance increases as the lag distance increases, indicating that samples which are located closer to each other are more similar than those farther away. The sill is the maximum of variance and the range is a lag distance at which the sill reaches constantly or asymptotically. Practically, the effective range is the distance where the function reaches 95% of its sill. The sill marks the variation level, while the range means the maximum spatial autocorrelation distance. Locations farther apart than the range are spatially uncorrelated or independent (Webster and Oliver, 2007).

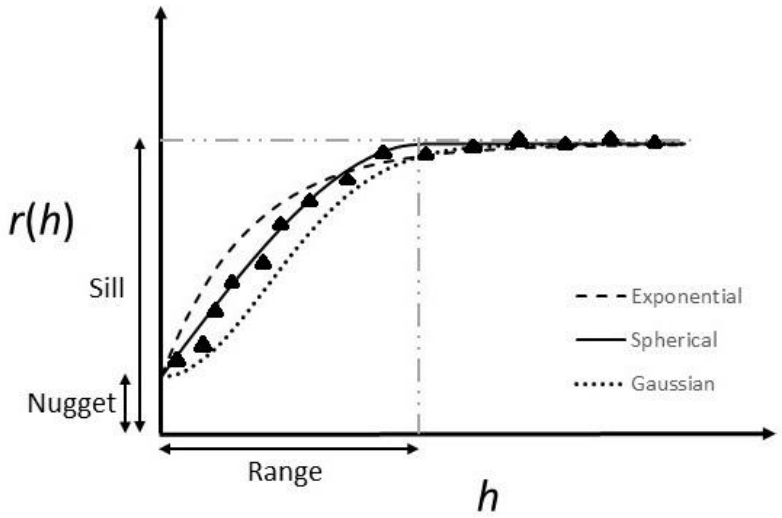


Figure 3.3: Bounded semivariogram models with the same range, reproduced from Goovaerts (1997).

(b) Semivariogram models

As we have mentioned that the experimental semivariogram is a graph of plotted points, we need model a function that follows the general trend, ignoring the point-to-point fluctuation. The reason is that it is hard to separate error and structural increasing clearly. After fitting a continuous function, we can predict values at unsampled locations (Webster and Oliver, 2007). Gaussian, exponential and spherical models are frequently used bounded models (Fig. 3.3). The spherical model has actual range, while the exponential and Gaussian models have effective ranges. The respective Equations (3.4), (3.5) and (3.6) are listed at below:

Gaussian model with effective range *a* and sill *c*:

$$r(h) = \mathrm{c}\left\{1 - \exp\left(-3\frac{h^2}{a^2}\right)\right\} \tag{3.4}$$

Exponential model with effective range *a* and sill *c*:

$$r(h) = \mathrm{c}\left\{1 - \exp\left(-3\frac{h}{a}\right)\right\} \quad (3.5)$$

Spherical model with range *a* and sill *c*:

$$r(h) = \begin{cases} c\left\{1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right\}, & h \le a \\ c, & h > a \end{cases} \quad (3.6)$$

3.3.4.2 Interpolation

Spatial interpolation is the estimation of the variable value for all the points within the map area based on the observed values. It is a key process to convert discrete data points into a continuous distribution. Spatial interpolation relies on two cornerstones. The first one is the continuous variable of the surface and the second one is spatial autocorrelation (Chou, 1997). The spatial autocorrelation is based on the first law of geography of Waldo Tobler. He pointed out that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). The distribution maps produced from different interpolation methods could be quite dissimilar.

(a) The IDW interpolation

The IDW is a relatively simple interpolation method, and it is often used to visualise the overall distribution pattern, especially with a large data set (Zhang, 2006). This method assumes that the closer points are assigned greater weights than farther places when predicting the target point value. The IDW is defined by the specific relationship between spatial weight and distance, which is an inverse function (e.g. a square inverse or a cubic inverse). The power value decides how fast the spatial weight decreases as the increase of distance. The number of neighbours determines how many data points are included to predict the target point value. These two parameters are the key for an IDW.

(b) Kriging

The prediction of kriging is also a weighted combination of the data with the sum of 1. The spatial weights of kriging are allocated to the sample data within the neighbourhood of the target point, which are based on spatial autocorrelation rather than a function of inverse distance. As a result, kriging can provide more accurate estimation compared to other methods of spatial interpolation. Prior to the implementation of kriging, a model must be chosen to fit the sample semivariogram (Webster and Oliver, 2007). The ordinary kriging is one of the most frequently used classes of the kriging family when the mean of the variable is not known. The OK is a robust and unbiased estimator since the error variance is minimized and the error mean is equal to zero (Goovaerts, 1997). In an OK, the weight depends on the selected model of semivariogram and the position of the observed points (Krige, 1966).

(c) Block kriging

Block kriging is a popular approach to transfer point data to areal data. The point kriging refers to the prediction on point support, while block kriging estimates the averages over a block. In the study, prior to carrying out block kriging, a fine grid (100 m × 100 m) was laid over the GLA area. Then OK was used to predict Pb concentration at each node of the grid. In the end, the predicted Pb concentrations within a ward were averaged to represent the Pb concentration of the ward (Young et al., 2009). Block IDW used in the study is alike with block kriging whose estimator is IDW interpolation rather than the OK.

(d) The measurements for the accuracy of the interpolation

Three statistical parameters based on the result of cross-validation are applied to compare the accuracy of the interpolation methods. The mean relative error (MRE) and the root mean square error (RMSE) are calculated according to Equation (3.7) and (3.8), respectively.

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{O_i - E_i}{O_i}\right| \tag{3.7}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[O_i - E_i]^2} \tag{3.8}$$

where O_i is the observed value at location i, E_i is the estimated value by interpolation method at location i, and n is the sample size. The smaller is the MRE and RMSE, the more precise does the interpolation perform.

The correlation coefficient (R^2) is another criterion which is employed to detect the correlation between the measured and predicted values (Yao et al., 2014). It can be calculated using the following Equation (3.9):

$$R^2 = \frac{\left[\sum_{i=1}^{n}(O_i - O_{ave})(E_i - E_{ave})\right]^2}{\sum_{i=1}^{n}(O_i - O_{ave})^2 \sum_{i=1}^{n}(E_i - E_{ave})^2} \tag{3.9}$$

where O_{ave} is the average value of the measured samples, E_{ave} is the average value of the predicted values.

3.3.4.3 Moran's I

Moran's I is a global indicator of spatial association/autocorrelation (GISA) to measure whether the values of labile P are clustered, dispersed or random by taking account of the spatial distribution as a whole. It is calculated by Equation (3.10) (Cliff and Ord, 1981):

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \times \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad (3.10)$$

where n is the sample number of the variable x, x_i and x_j are the values at the location i and j, $\bar{x}$ is the mean and w_{ij} is the weight matrix. The weight determines a local neighborhood surrounding the location i, and x_i is compared to the weighted average of the values of its neighbors.

Moran's I usually ranges from -1 (dispersed, negative spatial autocorrelation with different high and low value patterns in the distance band) to 1 (clustered, positive spatial autocorrelation with the same high value or low value patterns in the distance band), and a value close to zero represents absence of spatial autocorrelation (randomness). The Moran's I statistic is assessed by testing a null hypothesis that the spatial pattern is random. In this study, we rejected the null hypothesis and concluded that the spatial patterns were clustered, because the Z score for Moran's I was far more than 1.96 ($p < 0.05$). The results were based on 9999 permutations to avoid sensitivity on the specific randomization (Huo et al., 2012).

Correlogram is a graph of Moran's I plotted against the size (distance) of the local neighbourhood. It is often used to examine the average radius of a spatial patch in a data set. If the Moran's I fluctuates around zero for all distance lag, the data set is spatial randomness. If the Moran's I departs from zero significantly, the data set has spatial autocorrelation.

3.3.4.4 Local Moran's I

Moran's I has limit to recognize local or regional spatial autocorrelation of the observations. Local Moran's I belongs to LISA family and can well figure the limitation out. It decomposes the global indicator into sub-indicators of each unit, with its sum for all the units being proportional to the corresponding global one (Zhu et al., 2018). It is used to subgroup the

positive autocorrelation into high-high cluster (high values surrounded predominantly by high values) or low-low cluster (low values surrounded predominantly by low values) and the negative autocorrelation into high-low outlier (high values surrounded primarily by low values) or low-high outlier (low values surrounded primarily by high values). The local Moran's I at location i is given by Equation (3.11) (Anselin, 1995; Getis and Ord, 1996):

$$I_i = \frac{z_i - \overline{z}}{\delta^2} \sum_{j=1, j \neq i}^{n} \left[w_{ij} (z_j - \overline{z}) \right] \quad (3.11)$$

where z_i is the value of the variable z at location i; $\overline{z}$ is the mean value of z of the n samples; z_j is the value of the variable at all the other locations (where $j \neq i$); σ^2 is the variance of variable z; and w_{ij} is a spatial weight which can be defined by many means. For example, we chose the fixed distance band of 5000 m in London P study and 50 nearest neighbors in London Al, Ca, Pb study.

When the local Moran's I is tested as significance using random permutations of 9999 times, it detects clusters as either "hot spots" (high-high cluster; high values and low values are classified based on the mean value), or "cold spots" (low-low cluster), as well as "spatial outliers" as either high-low outlier or low-high outlier.

3.3.5 Computer software

Original data were stored in a file of MS Excel®. Summary statistical parameters were given by SPSS® ver. 23. The computation of local Moran's I index was carried out in the GeoDa (version 1.6.7 by Luc Anselin, 2015) software. All the maps were generated using ArcGIS (ver. 10.3) software. The semivariograms were illustrated in ArcGIS. The Moran's I was calculated using spatial autocorrelation tool of spatial statistics tools in ArcGIS. The spatial correlograms were drawn by free R program.

3.4 Summary

This chapter has described the background of the sampling sites, the GLA area of the UK and Lake Taihu of China. The sampling and laboratory procedures have been provided. The methodologies and data analysis have been listed with details.

Reference

Aczel, A. D., 1995. Statistics: concepts and applications. IRWIN, USA. ISBN 0-256-11935-X.

Allen, M. A., Cave, M. R., Chenery, S. R. N., Gowing, C. J. B., Reeder, S., 2011 Sample preparation and inorganic analysis for urban geochemical survey soil and sediment samples. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 28-46.

Anselin, L., 1995. Local indicators of spatial association - LISA. Geogr. Anal. 27, 93-115.

Appleton, J. D., Johnson, C. C., Ander, E. L., Flight, D. M. A., 2013. Geogenic signatures detectable in topsoils of urban and rural domains in the London region, UK, using parent material classified data. Appl. Geochem. 39, 169-180.

BGS (British Geological Survey), 2011. British Geological Survey Official Website. Available online at: http://www.bgs.ac.uk/gbase/londonearth.html/ (Accessed 31 May 18).

Chou, Y-H., 1997. Exploring spatial analysis in geographic information systems. OnWord Press, Santa Fe, USA. ISBN: 1-56690-119-7.

Clark, I., Harper, W. V., 2000. Practical geostatistics 2000. Ecosse North America LLC, Columbus, Ohio USA. ISBN: 0-9703317-0-3.

Cliff, A. D., Ord, J. K., 1981. Spatial Processes: Models and Applications. Pion Limited, London.

Cressie, N.A.C., 1993. Statistics for Spatial Data, revised Edition edn. Wiley, New York.

Ding, S., Chen, M., Gong, M., Fan, X., Qin, B., Xu, H., Gao, S., Jin, Z., Tsang, D. C. W., Zhang, C., 2018. Internal phosphorus loading from sediments causes seasonal nitrogen limitation for harmful algal blooms. Sci. Total Environ. 625, 872-884.

Ding, S. M., Jia, F., Xu, D., Sun, Q., Zhang, L., Fan, C., Zhang, C. S., 2011. High-resolution, two-dimensional measurement of dissolved reactive phosphorus in sediments using the diffusive gradients in thin films technique in combination with a routine procedure. Environ. Sci. Technol. 45(22), 9680-9686.

Ding, S., Wang, Y., Xu, D., Zhu, C., Zhang, C., 2013. Gel-based coloration technique for the submillimeter-scale imaging of labile phosphorus in sediments and soils with diffusive gradients in thin films. Environ. Sci. Technol. 47(14), 7821-7829.

Gallardo, A., Paramá, R., 2007. Spatial variability of soil in two pant communities of NW Spain. Geoderma 139, 199-208.

Gao, G., Zhu, G. W., Qin, B.Q., Chen, J., Wang, K., 2006. Alkaline phosphatase activity and the phosphorus mineralization rate of Lake Taihu. Sci. China Earth Sci. 149, 176-185.

Getis, A., Ord, J. K., 1996. Local spatial statistics: an overview. In: Longley P, Batty M, editors. Spatial Analysis: Modelling in a GIS Environment. Cambridge, England: Geoinformation International, pp. 261-277.

Goovaerts, P., 1997. Geostatistics for natural resources evaluation. Biomedware, Inc & PGeostat, LLC, Ann Arbor, Michigan. Oxford University Press, New York.

Groeneveld, R. A., Meeden, G., 1984. Measuring Skewness and Kurtosis. The Statistician 33(4), 391-399.

Huo, X. N., Li, H., Sun, D. F., Zhou, L. D., Li, B. G., 2012. Combining geostatistics with Moran's I analysis for mapping soil heavy metals in Beijing, China. Int. J. Environ. Res. Public Health 9(3), 995-1017.

Johnson, C. C., 2011. Understanding the quality of chemical data from the urban environment – Part 1: Quality control procedures. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 61-76.

Krige, D. G., 1966. Two dimensional weighted moving average trend surfaces for ore-evaluation. J. South Afr. Inst. Min. Metall. 66, 13-38.

Li, G., Sun, G. X., Ren, Y., Luo, X. S., Zhu, Y. G., 2018. Urban soil and human health: a review. Eur. J. Soil Sci. 69, 196-215.

Lin, C. C., Mudholkar, G. S., 1980. A simple test for normality against asymmetric alternatives. Biometrika 67(2), 455-461.

Lu, Q., Chang, N-B. Joyce, J., Chen, A. S., Savic, D. A., Djordjevic, S., Fu., G., 2018. Exploring the potential climate change impact on urban growth in London by a cellular automata-based Markov chain model. Computers, Environment and Urban Systems. 68, 121-132.

Miles, J. C. H., Appleton, J. D., 2005. Mapping variation in radon potential both between and within geological units. J. Radiol. Prot. 25, 257-276.

Niemietz, K., Wagner, A., Gründig-Wendrock, B., Stoyan, D., Niklas, J. R., 2010. Variogram analysis of charge-carrier effective lifetime topograms in mc-Si materials. Sol. Energy Mater. Sol. Cells, 94, 164-170.

Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46(Supplement), 234-240.

Wang, J., Jiang, X., Zheng, B., Chen, C., Kang, X., Zhang, C., Song, Z., Wang, K., Wang, W., Wang, S., 2016. Effect of algal bloom on phosphorus exchange at the sediment-water interface in Meiliang Bay of Taihu Lake, China. Environ. Earth Sci. 75, 57. https://doi.org/10.1007/s12665-015-4810-z

Webster, R., Oliver, M. A., 2007. Geostatistics for environmental scientists. John Wiley & Sons, Ltd. England. ISBN-13: 978-0-470-02858-2 (HB).

Widerlund, A., Nowell, G. M., Davison, W., Pearson, D. G., 2012. High-resolution measurements of sulphur isotope variations in sediment pore-waters by laser ablation multicollector inductively coupled plasma mass spectrometry. Chem. Geol. 291, 278-285.

Yao, L., Huo, Z., Feng, S., Mao, X., Kang, S., Chen, J., Xu, J., Steenhuis, T. S., 2014. Evaluation of spatial interpolation methods for groundwater level in an arid inland oasis, northwest China. Environ. Earth Sci. 71, 1911-1924.

Ye, R., Shan, K., Gao, H., Zhang, R., Wang, S., Qian, X., 2015. Long-term seasonal nutrient limiting patterns at Meiliang Bay in a large shallow and subtropical Lake Taihu, China. J. Limnol. 74(3), 519-529.

Young, L. J., Gotway, C. A., Yang, J., Kearney, G., DuClos, C., 2009. Linking health and environmental data in geographical analysis: It's so much more than centroids. Spatial Spatio-temporal Epidemiol. 1, 73-84.

Zhang, C., 2006. Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway, Ireland. Environ. Pollut. 142, 501-511.

Zhang, C., Fay, D., McGrath, D., Grennan, E., Carton, O. T., 2008. Use of trans-Gaussian kriging for national soil geochemical mapping in Ireland. Geochem-Explor. Env. A. 8, 255-265.

Zhu, B., Fu, Y., Liu, J., He, R., Zhang, N., Mao, Y., 2018. Detecting the priority areas for health work allocation with LISA functions: an empirical analysis for China. Heal. Services Res. 18, 957. https://doi.org/10.1186/s12913-018-3737-y

Chapter 4
Research Papers

4.1 Spatial distribution patterns of P in topsoils of the GLA area and their natural and anthropogenic factors

Meng, Y., Cave, M., Zhang, C., 2018. Spatial distribution patterns of phosphorus in topsoils of Greater London Authority area and their natural and anthropogenic factors. Applied Geochemistry 88, 213-220.

Summary:

This article presented the hot spots and cool spots of P in the GLA area and explored the natural and anthropogenic factors that might affect the spatial pattern. The hot spot map was produced using local Moran's I based on 6467 topsoils samples that were collected by the BGS. Both the hot spot map and the results of ANOVA showed a remarkably natural control of P. The elevated P concentrated in the PMs of alluvium and river terrace deposits. The absence of hot spots in the lower Thames Estuary indicated that the tide diluted the P-enriched sediments. The high concentration of Silicon (Si) and low pH level corresponded to the cool spots of P in Hyde Park and Richmond Park. Except from the natural control, the hot spots in the city centre and built-up area suggested that soil P content was heavily influenced by human activities. The result of a t-test displayed the significant difference of P concentrations between urban area and non-urbanised area, implying that urbanization and built-up materials contributed to the accumulation of P. In conclusion, both natural and anthropogenic factors had impacts on the spatial patterns of P in the GLA area.

My dedication in this paper accounted for 90% in reviewing literatures, exploring data and writing manuscript.

Applied Geochemistry 88 (2018) 213–220

Contents lists available at ScienceDirect

Applied Geochemistry

journal homepage: www.elsevier.com/locate/apgeochem

Spatial distribution patterns of phosphorus in top-soils of Greater London Authority area and their natural and anthropogenic factors

Yuting Meng [a], Mark Cave [b], Chaosheng Zhang [a, *]

[a] International Network for Environment and Health, School of Geography and Archaeology, National University of Ireland, Galway, Ireland
[b] British Geological Survey, Environmental Science Centre, Nottingham, United Kingdom

ARTICLE INFO

Article history:
Received 8 January 2017
Received in revised form
30 May 2017
Accepted 31 May 2017
Available online 3 June 2017

Keywords:
London
Phosphorus (P)
GIS
Local Moran's I
Urbanization
The Thames Estuary
Parent materials (PMs)

ABSTRACT

Soil phosphorus (P) has a strong impact on soil and water quality. Soils in urban areas tend to enrich P, however, they have not been adequately investigated. A total of 6467 top-soil samples were collected and analysed by the British Geological Survey, providing basic data for studying the top-soil P distribution patterns and their environmental implications. The hotspots and cool spots were identified using the index of local Moran's I, which is a powerful methodology for discerning spatial clusters and spatial outliers. Combined with the results of one-way analysis of variances (ANOVA), a strong natural control of P was illustrated with elevated concentrations in areas of alluvium and river terrace deposits. P concentration in the lower Thames Estuary was clearly influenced by the tidal effect which has diluted the P-enriched sediments. The high concentration of Si and low pH level were linked to the low value clusters of P in Hyde Park and Richmond Park. Besides the natural control, the high value clusters concentrated in the city centre and built-up area, which indicated soil P content was strongly affected by human activities. The results of a *t*-test also showed the significant distinction of P concentrations between urban area and non-urbanised area, implying that urbanization and built-up materials accounted mostly for the locations and magnitude of the P pool. To conclude, the spatial patterns of P observed in the study area were controlled by both natural (parent materials (PMs) and geomorphology) and anthropogenic (urbanization) factors.

1. Introduction

With the worldwide rapid urbanization, urban soils are receiving considerable attention due to their associations with the life quality of human beings (Norra and Stüben, 2003). Phosphorus (P), as a limiting nutrient for organisms, is responsible for water eutrophication (Conley et al., 2009; Halecki and Gąsiorek, 2015). The P-enriched urban soils can affect groundwater quality via leaching (Zhang et al., 2001). Moreover, P derived from urban household waste can decrease the stability of organic carbon (C) in urban soils, resulting in organic C loss (Chen et al., 2014). Therefore, the locations and scale of soil P accumulation provides crucial information for P management, in particular preventing P leaching into water bodies.

Previous researches have revealed that P enriched in the soils of a number of big cities, including Nanjing (Zhang et al., 2001), Hangzhou (Zhang, 2004), Beijing (Xia et al., 2013), Nanchang (Chen et al., 2014) of China, Bangkok (Faerage et al., 2001) of Thailand, Gälve (Nilsson, 1995) of Sweden and Phoenix (Metson et al., 2012) of the USA. There are a number of factors influencing P accumulation and sequestration in urban areas, such as parent materials (PMs), hydrology, biotic processes, and current and historical land use and management (Bennett, 2003; Bennett et al., 2004). These factors can be classified into two categories: the natural (internal) factors and the anthropogenic (external) factors. The natural source of P comes from PM which plays a role in driving differences in P status in long-timescale (Mage and Porder, 2013). Appleton et al. (2013) reported that PMs control 12–16% of the variance of P in the London region. The external sources of P are composed of fertilizer application for agricultural purpose in fields and green spaces, food and human waste, building materials like asphalt, wood and cement (Metson et al., 2012), among which sewage treatment works and septic tanks are two main inputs of P to urban soils (Neal et al., 2005). The phenomenon of P accumulation in surface soils implies that human activities can modify P concentration to some extent. Interestingly, P has long been used to

* Corresponding author.
E-mail address: Chaosheng.Zhang@nuigalway.ie (C. Zhang).

http://dx.doi.org/10.1016/j.apgeochem.2017.05.024
0883-2927/

indicate human activities in the field of archaeology during pre-agricultural and agricultural ages (Holliday and Gartner, 2007).

The study area, London, is the capital of the United Kingdom with a two-thousand-year history as a major settlement. The population was 8.2 million based on the 2011 Census and will continue to increase in the next few decades (Great London Authority, 2016). The land use and soil quality could be changed dramatically in suburban areas. The local government makes efforts to upgrade sewage (including sludge) treatment capacity and develop the Thames Tideway Sewer Tunnels in order to address the issues of sewer overflows and improve the water quality (Great London Authority, 2016). Consequently, to reduce economic cost and increase the efficiency of controlling P contamination, the hotspots which pose a potentially high risk should be targeted critically and the input sources should be assessed (Huang et al., 2012).

When considering a mixture of natural and anthropogenic controls, geographical information system (GIS) and spatial analyses, for example, inverse distance weighted (IDW) interpolation and local Moran's I, are powerful tools to explore P spatial distribution and identify of hotspots and cool spots. Hotspots are the area with high values of P concentration, oppositely, cool spots represent the area with low values. The local Moran's I, one of Local Indicators of Spatial Association (or Autocorrelation) (LISA) methods, has been widely used for extraction of spatial patterns (Bone et al., 2013; Voutchkova et al., 2014; Majumdar and Biswas, 2016). In terms of soil contamination researches, hotspots often represent contamination sites in comparison to the values of their neighbours (Zhang et al., 2008).

In this study, the GIS mapping techniques and spatial analyses are implemented to address the issues as follow: i) to reveal the spatial patterns of P in top-soils of the Greater London Authority (GLA) area; ii) to identify the locations of P accumulation and depletion; and iii) to explore the natural and anthropogenic factors causing the spatial patterns.

2. Materials and methods

2.1. Soil parent material

A surface PM map (Fig. 1) with simplified geological types (Miles and Appleton, 2005) was adopted for the study. Cretaceous and Palaeogene bedrocks cover the most of the GLA area, meanwhile some parts of the area are underlain by substantial Quaternary superficial deposits (British Geological Survey, 2011). The white chalk subgroup, formed in the Cretaceous, is the oldest bedrock within the GLA area, while the Thanet sand formation is the oldest Palaeogene deposit. The Lambeth group, Thames group, and Bracklesham group belong to Palaeogene deposits, cropping out extensively, among which Thames group is the most widespread. The Quaternary deposits include clay-with-flints, glacial till, river terrace deposits, alluvium, and Head. The clay-with-flints formed from the original Palaeogene cover, and earlier Quaternary deposits are heterogeneous in texture, colour, and clast content. Glacial till is deposited by the Anglian ice sheet which comprises mostly of stiff pebbles or boulders and silty clay. River terrace deposits have been formed by the diversion of the River Thames, and the gravel deposition crops out mainly on hilltops. Alluvial deposits that primarily form a flat surface in valleys of the rivers Thames and Lee are most recent deposits within the research area. Generally, the alluvium consists principally of silty clay and fine-to coarse-grained sands.

2.2. Soil chemistry data

In the GLA area, a hand-held auger was used to collect 6467 top-soil samples by British Geology Survey. The sampling density was 4 samples per square kilometre and the sampling depth was ca. 5–20 cm. At each site, a composite sample based on 5 sub-samples was taken at the centre and four corners of a 20 m square. Soil samples were air-dried and sieved through a nylon sieve with 2 mm apertures. Coarse powder were sealed into vessels containing mill balls after simply disaggregated by hand in a pestle and mortar. Milled powder was compressed into pellets with the addition of solid binder which was used to improve adhesion of urban samples (Johnson, 2011). Forty-eight trace and major chemical elements were measured by X-ray fluorescence (XRF) spectrometry. Besides, loss on ignition (LOI at 450 °C) and pH were also determined. Details of sample preparation, analytical methods, and quality control procedures were described in Allen et al. (2011) and Johnson (2011).

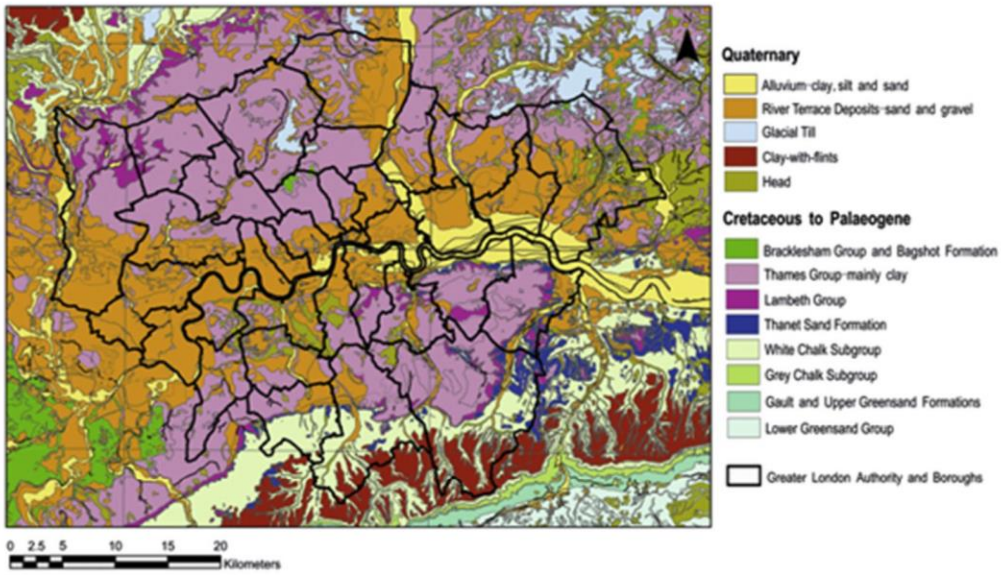


Fig. 1. Geology map of the London region.

2.3. Statistical and geostatistical analyses

The one-way analysis of variance (ANOVA) was applied to determine the variation between PMs. Normal probability plot suggested that the data distribution needed to be normalized prior to ANOVA analysis. Therefore, log transformation was performed to improve the normality to meet the statistical assumption. Subsequently, Tukey's multiple comparisons method was employed to further observe differences among distinct PMs.

Geostatistical analysis and GIS mapping techniques were used to present spatial distribution maps and to identify the hotspots, in order to explore the probable sources of P. The relatively simple interpolation method of IDW is sufficient to visualise the overall contamination pattern (Zhang, 2006). IDW is based on the assumption that each existing point datum has a local impact that decreases with distance. There are two main parameters for an IDW. The first one is the power value, which decides how rapidly the influence of the point falls off with distance. The higher the power is, the more influence the nearest neighbour has to the prediction location. The second one is the number of neighbours to be included which is to some extent arbitrary (Zhang et al., 2011).

The P concentration hotspots can be recognized using the local Moran's I index (Anselin, 1995; Getis and Ord, 1996):

$$I_i = \frac{z_i - \bar{z}}{\delta^2} \sum_{j=1, j\neq i}^{n} [w_{ij}(z_j - \bar{z})]$$

where z_i is the value of the variable z at location i; $\bar{z}$ is the mean value of z of the n samples; z_j is the value of the variable at all the other locations (where $j \neq i$); σ^2 is the variance of variable z; and w_{ij} is a weight which can be defined in different ways, such as a fixed distance band (Zhang et al., 2008). The weight w_{ij} can also be generated according to the specific spatial relationships among all the features in a dataset.

Spatial clusters, including high-high clusters and low-low clusters, are indicated by high positive local Moran's I values, which have similarly high or low values as their neighbours. Spatial outliers, containing high-low clusters and low-high clusters, are identified by high negative local Moran's I values, which behave differently from their neighbours (Lalor and Zhang, 2001). The significance level of Local Moran's I index can be assessed on an assumption of a normal distribution due to standardization (Anselin, 1995). In this study, the significance level was set at 0.05 and the resulted map was based on 999 permutations to avoid sensitivity on the specific randomization. More detailed explanations of local Moran's I are available in Anselin (1995) and Zhang et al. (2008).

2.4. Data analyses and computer software

Raw data were stored in a file of MS Excel®. Basic statistical parameters were obtained and the test for normality was carried out using SPSS® ver. 23. The calculation of Local Moran's I index was performed in the software GeoDa (version 1.6.7 by Luc Anselin, 2015). All maps were produced using ArcGIS (version 10.3) software.

3. Results and discussion

3.1. Natural factors

3.1.1. Influences of PMs

Summary statistics (Table 1) demonstrate that P concentrations had a strong positively skewed distribution (Fig. 2). The median value of P concentration in the surface soils in GLA area was 1397 mg/kg, which was markedly higher than European agriculture soil median value of 785 mg/kg (Reimann et al., 2012).

Fig. 3(a) shows that the majority of hotspots overlap two typical PMs, alluvium and river terrace deposits, from central-west and central-north to city centre. The elevated P concentrations are associated with alluvium and river terrace deposit due to the impact of hydrology, erosion processes and transformations of various forms of P (Hunter and Walton, 2008; Neal et al., 2005). Most alluvium is carried and deposited during floods when erosion is most active and the carrying capacity is at a maximum (Turowski et al., 2013). Alluvial deposits which are the most recent deposits within the GLA area form an approximately flat surface in valley floors and occur mainly in the valleys of the River Thames and River Lee. On average, the alluvium consists largely of clayey silt and silty clay with subsidiary sands. Within the GLA area, river terrace gravels were deposited throughout the Thames valley, extending north into the Lee valley (British Geological Survey, 2011). Fluvial sediments are recognized as being extremely crucial in controlling P flux in water-sediment system (Haggard et al., 2007). The close association between hydrology and P was also implied by water area percentages map (Fig. 3(b)) based on land-use data (Gov.uk, 2005). The land-use types (Generalised Land Use Database) 2005 (Enhanced Basemap) contain domestic buildings, non-domestic buildings, gardens, green space, road, rail, paths, and water areas. Cooper et al. (1995) demonstrated that riparian soils tended to enrich a great quantity of readily-transportable forms of P. Besides, studies have shown that the riparian soils perform as transient

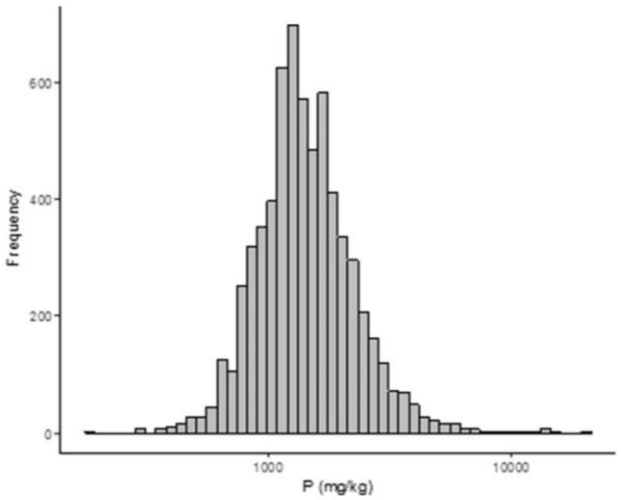


Fig. 2. Histogram of P for top-soil samples from London. Note that the P concentration is drawn in a log scale.

Table 1
Summary statistics of P (mg/kg) for top-soil samples from the London urban area. TN: total number of the samples; Min: minimum; Max: maximum; DL: detection limit.

TN	Min	5%	10%	25%	50%	75%	90%	95%	Max	Skew	DL
6467	175	698	829	1047	1397	1833	2488	3042	19595	5.58	218

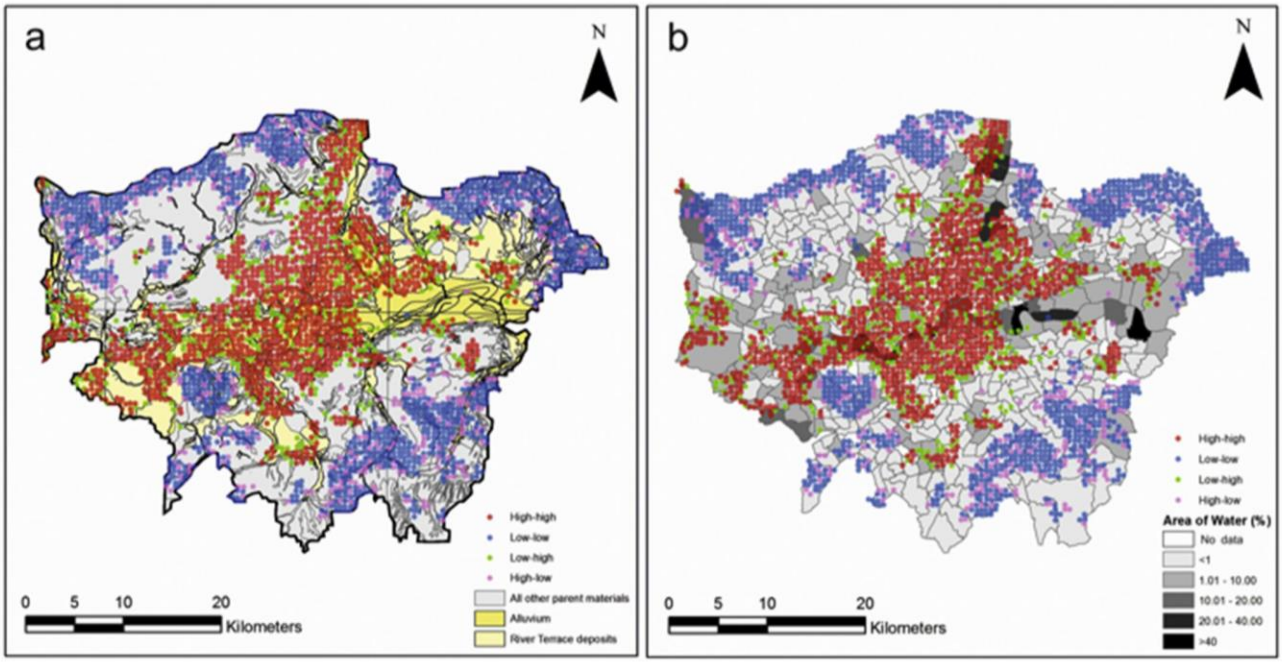


Fig. 3. (a) Spatial distribution map of significant hotspots and cool spots overlaying on geology map; (b) Spatial distribution map of significant hotspots and cool spots overlaying on water area percentages map.

storage of P between the aquifer and the stream which is in direct proportion to the P concentrations in streams (Froelich, 1988; Bridgham et al., 2001; Heeren et al., 2011). The sediments of adjacent water areas have a controlling function of retention and release of P between adjacent aquatic systems and uplands (Thompson and McFarland, 2010; Xia et al., 2011).

To further explore the association with fluvial sediment, one-way ANOVA was applied to examine whether there were significant differences among PMs. Log transformation was carried out to reach a relatively normal distribution, though transformed data still did not pass the Kolmogorov–Smirnov test (K–S test). Geochemical data rarely follow a log normal distribution, especially with large sample sizes (>1000) (Zhang et al., 2005). The result of one-way ANOVA showed a significant difference among PMs ($p < 0.001$). Given that Tukey HSD is helpful to determine specifically which type of PMs is different from each other, the results reveal that river terrace deposits, brickearth and river terrace deposit have higher concentrations of P than other types of PMs (Fig. 4). Based on the comprehensive results, the variability of P concentrations in top-soils of the GLA area is obviously linked to alluvium and river terrace deposits.

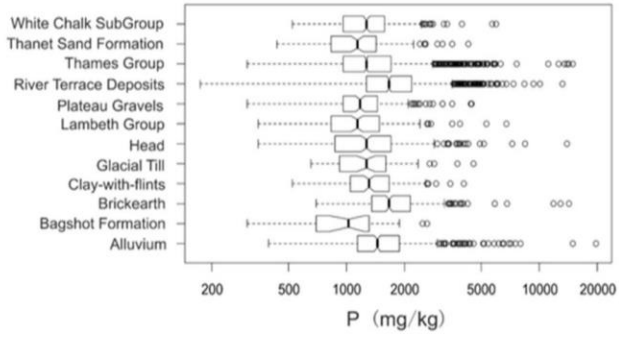


Fig. 4. Notched boxplot of P in top-soils of different PMs from London (n = 6467). Note that the P concentration is drawn in a log scale.

3.1.2. Influence of the Thames Estuary

Due to the close association between P concentrations with fluvial deposits, hotspots existed along the River Thames, the River Lee and Grand Union Canal, except for downstream below Tower Bridge of the Thames (Fig. 5) where high-high spatial clusters should be expected. This "special" pattern could be related to the Thames estuary, a typical macro tidal funnel-shaped estuary, where river flow, tides, and storm surges come into interaction (Uncles and Mitchell, 2011). According to Mikahailo and Mikahailova, (2012), the River Thames is divided into three zones. Firstly, from Teddington to near Tower Bridge, the suspended particulate matter (SPM) concentrations are low and mainly derived from inland. In this zone, the tidal effect is minimal and the bank sediment is mostly made up of the clay fraction. Secondly, from Woolwich reach to Gravesend reach, high SPM and a high rate of deposition exists. Downstream from the Tower Bridge are more influential by tides and the sediment is much coarser than that of the upstream. Lastly, below Woolwich reach and down to Southend, the sedimentation is gradually predominated by marine-derived bedload. The Tanshui Estuary, situated on the outskirts of the largest city of Taiwan, Taipei, showed a similar deposition pattern of P. The majority of total P derived from the upstream as well as city sewages deposited in the upper estuary and the rest was exported to the coast (Fang, 2000). Together with tides, drying and rewetting events leading P loss, the sediments of lower estuary and bay exhibit lower quantities of organic matter and P (Berbel et al., 2015; Gao et al., 2016).

3.1.3. Influences of silicon (Si) and pH

IDW is a quick deterministic interpolator without any assumptions which can be easily applied to produce an interpolated map.

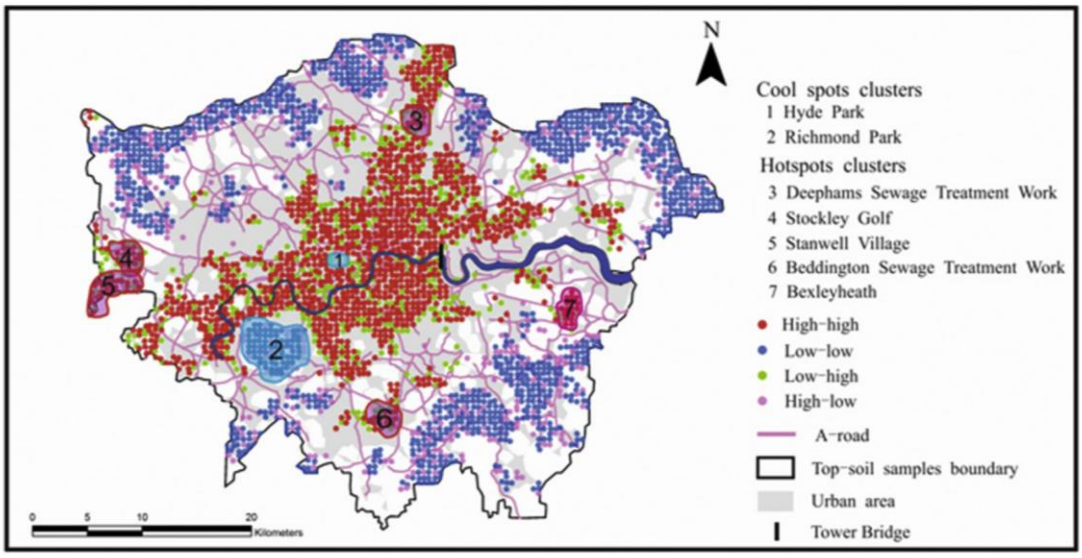


Fig. 5. Map of significant hotspots and cool spots of P overlaid onto a London street map.

In this study, the power value was set as 1 in order to reduce the "bulls-eye" effect around data locations; the number of neighbours was set as 16. The distribution maps for Si and pH values were shown in Fig. 6. London city centre had a high concentration of P, except for Hyde Park area (Fig. 5). Hyde Park, along with Kensington Gardens, Green Park, and St. James's Park constitute a continuous "green lung" in the heart of London. The most likely reason behind the relatively low concentrations in these parks was the less human activities in this area. Richmond Park was revealed as a significant cool spot in the map which is the largest royal park in London with nationally and internationally importance for wildlife conservation. Richmond Park is an upland where the human disturbance is rare and the concentration of Si is high in sandy soils (Fig. 6(a)). The Si content rises with the increase of sandstone in sandy soils, and decreases by dilution with organic matter and $CaCO_3$ (Bear, 1964). The Si distribution map shows that Richmond Park had high Si concentration (Fig. 6(a)) which indirectly meant low organic matter content. Given that nearly half the soil P occurs in combination with organic matter in surface soils, the P of this area was largely mineral or inorganic combination forms. As weathering proceeded in soils, the phosphates became increasingly bonded to aluminium (Al) and iron (Fe) rather than silicate minerals and are released from the sandy soils (Bear, 1964). Consequently, the soil containing the high concentration of Si was characteristic with low concentrations of P and other elements (Reimann et al., 2012). Moreover, as calcium phosphate in high pH alkaline and calcareous soils, the low pH of Richmond Park (Fig. 6(b)) was indicative of low inorganic P. With the development of acidity in soils, calcium phosphate is dissolved and precipitated as Al and Fe phosphate in leaching channels (Halvorson et al., 2016). In summary, high Si concentration and low level of pH synergistically foster the cool spots of P.

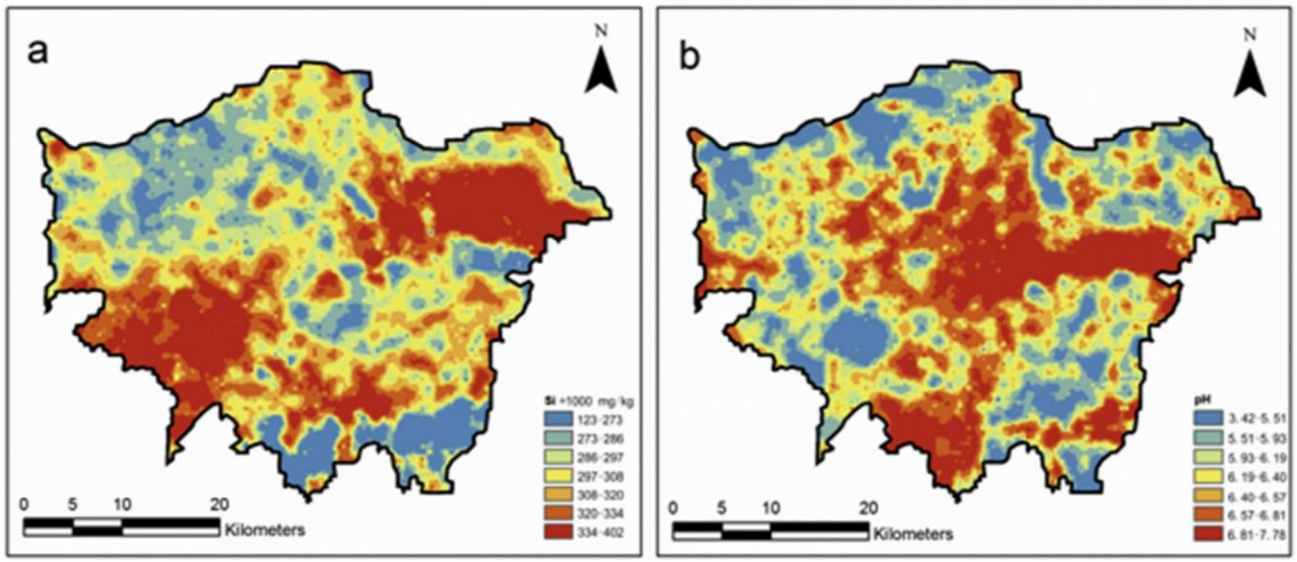


Fig. 6. (a) Spatial distribution map of Si; (b) Spatial distribution map of pH.

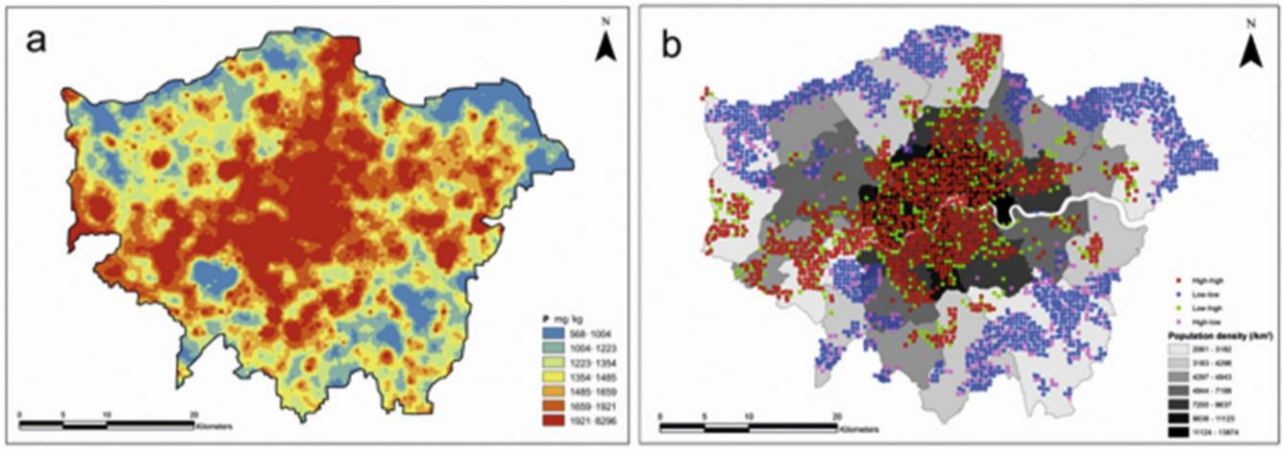


Fig. 7. (a) IDW spatial interpolation map of P; (b) Spatial distribution map of significant hotspots and cool spots overlay on the London population density map.

3.2. Anthropogenic factors

3.2.1. Influences of urbanization

The P distribution surface of the London urban area was created using IDW with raw data (Fig. 7(a)), displaying a pattern of high P values concentrated in the city centre as well as low values in suburban areas (Fig. 5). To find out the significance of urbanization, the data were divided into two sectors: the urban area and non-urbanised area. The urban area polygon was generated using the built-up tool in ArcGIS, on the basis of the building shapefile of the GLA area downloaded from Geofabrik (2016). The sample sizes were 3683 and 2784, and the median values were 1528 mg/kg and 1222 mg/kg in the urban area and non-urbanised area, respectively. The independent *t*-test was used to compare the difference in P concentrations between urban area and non-urbanised area after log transformation. Results from Levene's test suggested that the log transformed data set did not meet with the equal variances assumption ($p < 0.001$). Thus, the *t*-value of 18.919 with "equal variances not assumed" was used. There was a significant difference between urban area and non-urbanised area ($p < 0.001$) for P concentration (Fig. 8). The distribution pattern corresponded to the population density of London (Fig. 7(b)) which was produced according to London Census 2011 data (London datastore, 2011). Moreover, the median P concentration of London urban area was notably higher than that of European agricultural, grazing land (Reimann et al., 2012) and rural areas in England and Wales (Rawlins et al., 2012). Compared with Doncaster, Scunthorpe, Mansfield, and Sheffield, London still had slightly higher P concentration. (Table 2).

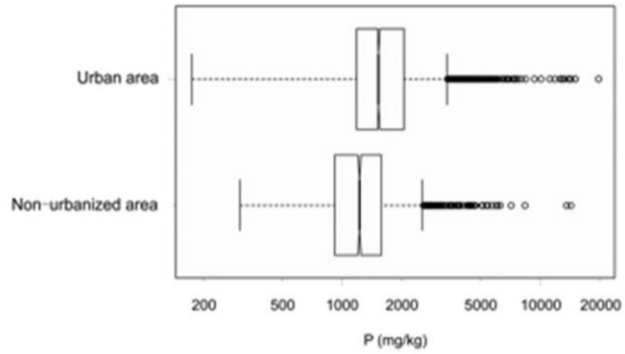


Fig. 8. Notched boxplot of P in top-soils of London urban area (n = 3683) and non-urbanized area (n = 2784). Note that the P concentration is drawn in a log scale.

There are a few aspects related to urbanization affecting the P behaviour in urban soils. As human waste and food count for the major parts of imported P in urban ecosystems, population density shapes the concentration pattern of P (Brett et al., 2005; Steinke et al., 2013). In addition, built-environment materials, such as asphalt, cement, and wood are contributing to the elevated P (e.g. the urbanized-area pattern is visible in Fig. 5). As a result, the P accumulation is associated with the levels of population and urbanization of the city. It is dangerous that the accumulated P could be mobilised into stream water by urban stormwater (Brezonik and Stadelmann, 2002), causing eutrophication of aquatic ecosystems (Yuan et al., 2007).

Table 2
Comparison of median P (mg/kg) in London urban top-soils with other urban, rural soils in England and Wales, European agricultural and grazing land soils. Min: minimum; Max: maximum; N: number of the samples; NA: not available.

	Min	5%	10%	25%	50%	75%	90%	95%	Max	N	Reference
London	175	698	829	1047	1397	1833	2488	3055	19595	6467	
Doncaster	131	524	698	829	1135	1484	2226	2793	5237	279	O' Donnell (2005a)
Lincoln	393	611	786	1047	1397	1964	2662	3841	5717	216	O' Donnell (2005b)
Mansfield	218	524	698	873	1178	1615	2357	2924	5281	257	Freestone et al. (2004a)
Scunthorpe	131	611	698	829	1135	1746	2357	2793	5281	196	O' Donnell (2005c)
Sheffield	218	611	786	1004	1353	1789	2357	2880	6590	575	Freestone et al. (2004b)
European agricultural soil	61	349	432	576	786	1104	1493	1872	4408	2108	Reimann et al. (2012)
European grazing land soil	87	305	385	532	777	1122	1532	1881	8772	2024	Reimann et al. (2012)
Rural areas in England and Wales	44	NA	480	650	870	1100	1400	NA	6100	5670	Rawlins et al. (2012)

3.2.2. Influences of sewage treatment works (STWs)

Apart from the large high-high value clusters concentrated within the city centre, several small hotspots were scattered in the outskirts, illustrating the spatial heterogeneity of urban soils (Fig. 5). The hotspot analysis showed that the areas around two sewage treatment works (STWs), Beddington and Deephams, had high P concentrations. In the UK, the major concern of P contamination is associated with sewage effluent sources from urban and farming (Vaze and Chiew, 2004; Neal et al., 2010b). Considerable money has been spent on dealing with P removal from final effluents at sewage works (Neal et al., 2010a). In 2013, the City of London wastewater treatment plants removed 91% of the P, however, flows exceeded the capacity of the treatment plants during rain events (City of London Environment and Engineering Services, 2014). Additionally, the majority of removed P becomes sludge. Although 75% of treated sludge is transferred to agricultural land (Shepherd et al., 2016) in England and Wales, the sludge residue remains as secondary contamination source of P which needs to be monitored. Moreover, there is a trunk sewer near Deephams sewage treatment work (Fig. 5) which may be another source of P input. Another finding is Stanwell Village (Fig. 5), located southwest of Heathrow Airport, where the high P concentration is possibly sewages related.

3.2.3. Influences of fertilizer application

Bexleyheath and Stockley Golf Course showed high levels of P (Fig. 5). Bexleyheath Golf Course is situated in the lowland, whilst Stockley Park Golf Course covers a much larger area, extending over 240 acres. Fertilizers are intensively added in golf courses (Shuman et al., 2000). P generally improves the growth of ornamental plants and turfgrass (McDowell et al., 2001) and is essential to the start-up or green-up phase (King and Balogh, 2011). Application of high concentrations of P fertiliser can lead to losses via sub-surface drains such as creeks and ponds in the golf course, thus degrading water quality (Shuman, 2003). Two sports fields of Townley Grammar School, close to Bexleyheath Golf Course, might also contribute to the hotspots of P in Bexleyheath area. Low dose applications, organic formulation, and reduced rate of fertilizer are recommended to improve P fertility management (King et al., 2012).

4. Conclusions

This study has demonstrated that the variation of P in top-soils of the GLA area was influenced by a combination of natural and anthropogenic controls. Natural factors, consisting of PM, the hydrology in the Thames Estuary, the distribution of Si and pH level in top-soils, were clearly associated with the P distribution. In addition, anthropogenic factors, such as population density, STWs and fertilizer application, also had an effect on the P concentration and spatial distribution, causing localised high concentrations. The combination of these effects and their interactions contributes to the unique P distribution patterns in the GLA area and the identified P hotspots. The spatial distribution patterns provide important information for soil management.

Acknowledgements

This study is sponsored by the Royal Society International Exchange Scheme of the UK (2015–2017, No. IE141447). The soil data used in this study were provided by the British Geological Survey from the London Earth Project.

References

Allen, M.A., Cave, M.R., Chenery, S.R.N., Gowing, C.J.B., Reeder, S., 2011. Sample preparation and inorganic analysis for urban geochemical survey soil and sediment samples. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 28–46.

Anselin, L., 1995. Local indicators of spatial association - LISA. Geogr. Anal. 27, 93–115.

Appleton, J.D., Johnson, C.C., Ander, E.L., Flight, D.M.A., 2013. Geogenic signatures detectable in topsoils of urban and rural domains in the London region, UK, using parent material classified data. Appl. Geochem 39, 169–180.

Bear, F.E., 1964. Chemistry of the Soil, second ed. Reinhold Publishing Corporation, New York, Amsterdam, London, pp. 126–133.

Bennett, E.M., 2003. Soil phosphorus concentrations in Dane County, Wisconsin, USA: an evaluation of urban-rural gradient paradigm. Environ. Manage 32, 476–487.

Bennett, E.M., Carpenter, S.R., Clayton, M.K., 2004. Soil phosphorus variability: scale-dependence in an urbanizing agricultural landscape. Landsc. Ecol. 20, 389–400.

Berbel, G.B.B., Favaro, D.I.T., Braga, E.S., 2015. Impact of harbour, industry and sewage on the phosphorus geochemistry of a subtropical estuary in Brazil. Mar. Poll. Bull. 93, 44–52.

Bone, C., Wulder, M.A., White, J.C., Robertson, C., Nelson, T.A., 2013. A GIS-based risk rating of forest insect outbreaks using aerial overview surveys and the local Moran's I statistic. Appl. Geogr. 40, 161–170.

Brett, M.T., Arhonditsis, G.B., Mueller, S.E., 2005. Non-point-source impacts on stream nutrient concentrations along a forest to urban gradient. Environ. Manage 35, 330–342.

Brezonik, P., Stadelmann, T.H., 2002. Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolitan area, Minnesota, USA. Water Res. 36, 1743–1757.

Bridgham, S.D., Johnston, C.A., Schubauer-Berigan, J.P., Weishampel, P., 2001. Phosphorus sorption dynamics in soils and coupling with surface and pore water in riverine wetlands. Soil Sci. Soc. Am. J. 65, 577–588.

British Geological Survey, 2011. British geological survey official website. Available online at: http://www.bgs.ac.uk/gbase/londonearth.html/ (Accessed 6 December 2016).

Chen, F.-S., Yavitt, J., Hu, X.-F., 2014. Phosphorus enrichment helps increase soil carbon mineralization in vegetation along an urban-to-rural gradient, Nanchang, China. Appl. Soil Ecol. 75, 181–188.

City of London Environment and Engineering Services, 2014. Thames river water quality 2013. Available online at: https://www.london.ca/residents/Environment/Rivers-Creeks/Documents/2013-Thames-River-Water-Quality-AODA.pdf/ (Accessed 6 December 2016).

Conley, D.J., Paerl, H.W., Howarth, R.W., Boesch, D.F., Seitzinger, S.P., Havens, K.E., Lancelot, C., Likens, G.E., 2009. Controlling eutrophication: nitrogen and phosphorus. Science 323, 1014–1015.

Cooper, A.B., Smith, C.M., Smith, M.J., 1995. Effects of riparian set-aside on soil characteristics in an agricultural landscape: implications for nutrient transport and retention. Agric. Ecosyst. Environ. 55, 61–67.

Faerage, J., Magid, J., Penning de Vries, F., 2001. Urban nutrient balance for Bangkok. Ecol. Model 139, 63–74.

Fang, T.H., 2000. Partitioning and behaviour of different forms of phosphorus in the Tanshui Estuary and one of its tributaries, northern Taiwan. Estuar. Coast. Shelf Sci. 50, 689–701.

Freestone, S.E., O'Donnell, K.E., Brown, S.E., 2004a. Geochemical Baseline Data for the Urban Area of Mansfield. British Geological Survey Internal Report, IR/02/082, p. 16.

Freestone, S.E., O'Donnell, K.E., Brown, S.E., 2004b. Geochemical Baseline Data for the Urban Area of Sheffield. British Geological Survey Internal Report, IR/02/084, p. 14.

Froelich, P.N., 1988. Kinetic control of dissolved phosphate in natural rivers and estuaries: a primer on the phosphate buffer mechanism. Limnol. Oceanogr. 33, 649–668.

Gao, Z., Fang, H., Bai, J., Jia, J., Lu, Q., Wang, J., Chen, B., 2016. Spatial and seasonal distributions of soil phosphorus in a short-term flooding wetland of the Yellow River Estuary, China. Ecol. Inf. 31, 83–90.

Geofabrik, 2016. Geofabrik official website. Available online at: http://download.geofabrik.de/europe/great-britain/england/greater-london.html/ (Accessed 6 December 2016).

Getis, A., Ord, J.K., 1996. Local spatial statistics: an overview. In: Longley, P., Batty, M. (Eds.), Spatial Analysis: Modelling in a GIS Environment. GeoInformation International, Cambridge, England, pp. 261–277.

Govuk, 2005. Gov. uk official website. Available online at: http://www.communities.gov.uk/documents/planningandbuilding/xls/154329.xls/ (Accessed 6 December 2016).

Great London Authority, 2016. The London Plan: the spatial development strategy for London consolidated with alternations since 2011. London, England. Available online at: www.london.gov.uk/what-we-do/planning/london-plan/current-london-plan (Accessed 6 December 2016).

Haggard, B.E., Smith, D.R., Brye, K.R., 2007. Variations in stream water and sediment phosphorus among select Ozark catchments. J. Environ. Qual. 36, 1725–1734.

Halecki, W., Gąsiorek, M., 2015. Seasonal variability of microbial biomass phosphorus in urban soils. Sci. Total Environ. 502, 42–47.

Halvorson, J.J., Schmidt, M.A., Hagerman, A.E., Gonzalez, J.M., Liebig, M.A., 2016. Reduction of soluble nitrogen and mobilization of plant nutrients in soils from U.S northern Great Plains agroecosystems by phenolic compounds. Soil Biol.

Biochem. 94, 211–221.
Heeren, D.M., Fox, G.A., Miller, R.B., Storm, D.E., Fox, A.K., Penn, C.J., Halihan, T., Mittelstet, A.R., 2011. Stage-dependent transient storage of phosphorus in alluvial floodplains. Hydrol. Process 25, 3230–3243.
Holliday, V.T., Gartner, W.G., 2007. Methods of soil P analysis in archaeology. J. Archaeol. Sci. 34, 301–333.
Huang, L., Wang, H., Li, Y., Lu, S., 2012. Spatial distribution and risk assessment of phosphorus loss potential in urban-suburban soil of Lishui, China. Catena 100, 42–49.
Hunter, H.M., Walton, R.S., 2008. Land-use effects on fluxes of suspended sediments, nitrogen and phosphorus from a river catchment of the Great Barrier Reef, Australia. J. Hydrol. 356, 131–146.
Johnson, C.C., 2011. Understanding the quality of chemical data from the urban environment – Part 1: quality control procedures. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 61–76.
King, K.W., Balogh, J.C., 2011. Stream water nutrient enrichment in a mixed-use watershed. J. Environ. Monit. 13, 721–731.
King, K.W., Balogh, J.C., Agrawal, S.G., Tritabaugh, C.J., Ryan, J.A., 2012. Phosphorus concentration and loading reductions following changes in fertilizer application and formulation on managed turf. J. Environ. Monit. 14, 2929–2938.
Lalor, G., Zhang, C.S., 2001. Multivariate outlier detection and remediation in geochemical databases. Sci. Total Environ. 281, 99–109.
London datastore, 2011. London Datastore official website. Available online at: http://data.london.gov.uk/census/data/ (Accessed 6 December 2016).
Mage, S.M., Porder, S., 2013. Parent material and topography determine soil Phosphorus status in the Luquillo Mountains of Puerto Rico. Ecosystems 16, 284–294.
Majumdar, D.D., Biswas, A., 2016. Quantifying land surface temperature change from LISA clusters: an alternative approach to identifying urban land use transformation. Landsc. Urban Plan 153, 51–65.
McDowell, R.W., Sharpley, A.N., Condron, L.M., Haygarth, P.M., Brookes, P.C., 2001. Processes controlling soil phosphorus release to runoff and implications for agricultural management. Nutr. Cycl. Agroecosys 59, 269–284.
Metson, G.S., Hale, R., Iwaniec, D.M., Cook, E.M., Corman, J.R., Galletti, C.S., Childers, A.L., 2012. Phosphorus in Phoenix: a budget and spatial representation of phosphorus in an urban ecosystems. Ecol. Appl. 22, 705–721.
Mikahailo, V.N., Mikahailova, M.V., 2012. Tides and storm surges in the Thames river estuary. Water Resour. 39, 351–365.
Miles, J.C.H., Appleton, J.D., 2005. Mapping variation in radon potential both between and within geological units. J. Radiol. Prot. 25, 257–276.
Neal, C., House, W.A., Jarviea, H.P., Neal, M., Hill, L., Wickham, H., 2005. Phosphorus concentrations in the river Dun, the Kennet and Avon canal and the river Kennet, southern England. Sci. Total Environ. 344, 107–128.
Neal, C., Jarvie, H.P., Williams, R.J., Love, A., Neal, M., Wickham, H., Harman, S., Armstrong, L., 2010a. Declines in phosphorus concentration in the upper River Thames (UK): links to sewage effluent cleanup and extended end member mixing analysis. Sci. Total Environ. 408, 1315–1330.
Neal, C., Jarvie, H.P., Withers, P.J.A., Whitton, B.A., Neal, M., 2010b. The strategic significance of wastewater sources to pollutant phosphorus levels in English rivers and to environmental management for rural, agricultural and urban catchments. Sci. Total Environ. 408, 1485–1500.
Nilsson, J., 1995. A phosphorus budget for a Swedish municipality. J. Environ. Manage 45, 243–253.
Norra, S., Stüben, D., 2003. Urban soils. J. Soils Sediments 3, 230–233.
O' Donnell, K.E., 2005a. Geochemical Baseline Data for the Urban Area of Doncaster. British Geological Survey Internal Report IR/02/079, p. 13.
O' Donnell, K.E., 2005b. Geochemical Baseline Data for the Urban Area of Lincoln. British Geological Survey Internal Report IR/02/081, p. 14.
O' Donnell, K.E., 2005c. Geochemical Baseline Data for the Urban Area of Scunthorpe. British Geological Survey Internal Report IR/02/083, p. 16.
Reimann, C., Filzmoser, P., Fabian, K., Hron, K., Birke, M., Demetriades, A., Dinelli, E., Ladenberger, A., The GEMAS Project Team, 2012. The concept of compositional data analysis in practice – total major element concentrations in agricultural and grazing land soils of Europe. Sci. Total Environ. 426, 196–210.
Rawlins, B.G., McGrath, S.P., Scheib, A.J., Breward, N., Cave, M., Lister, T.R., Ingham, M., Gowing, C., Carter, S., 2012. The advanced soil geochemical atlas of England and Wales. British geological survey, keyworth, Nottingham, pp. 134–135. Available online at: www.bgs.ac.uk/gbase/advsoilatlasEW.html/ (Accessed 6 December 2016).
Shepherd, J.G., Sohi, S.P., Heal, K.V., 2016. Optimising the recovery and re-use of phosphorus from wastewater effluent for sustainable fertiliser development. Water Res. 94, 155–165.
Shuman, L.M., 2003. Fertilizer source effects on phosphate and nitrate leaching through simulated golf greens. Environ. Pollut. 125, 413–421.
Shuman, L.M., Smith, A.E., Bridges, D.C., 2000. Potential movement of nutrients and pesticides following application to golf courses. Fate and management of turfgrass chemicals. In: Clark, J.M., Kenna, M.P. (Eds.), ACS Symposium Series, vol. 743, pp. 78–93.
Steinke, K., Kussow, W.R., Stier, J.C., 2013. Potential contributions of mature prairie and turfgass to phosphorus in urban runoff. J. Environ. Qual. 42, 1176–1184.
Thompson, C.A., McFarland, A.M.S., 2010. Effects of surface and groundwater interactions on phosphorus transport within streambank sediments. J. Environ. Qual. 39, 548–557.
Turowski, J.M., Badoux, A., Leuzinger, J., Hegglin, R., 2013. Large floods, alluvial overprint, and bedrock erosion. Earth Surf. Process. Landf. 38, 947–958.
Uncles, R.J., Mitchell, S.B., 2011. Turbidity in the Thames Estuary: how turbid do we expect it to be? Hydrobiologia 672, 91–103.
Vaze, J., Chiew, F., 2004. Nutrient loads associated with different sediment sizes in urban stormwater and surface pollutants. J. Environ. Eng.-ASCE 130, 391–396.
Voutchkova, D.D., Emstsen, V., Hansen, B., Sørensen, B.L., 2014. Assessment of spatial variation in drinking water iodine and its implications for dietary intake: a new conceptual model for Denmark. Sci. Total Environ. 493, 432–444.
Xia, J.G., Zhong, Y.M., Cao, X.X., 2011. Relation between phosphorous release and soil character with alternative dry-wet condition. J. Soil Water Conserv. 25, 237–243.
Xia, X., Zhao, X., Lai, Y., 2013. Levels and distribution of total nitrogen and total phosphorus in urban soils of Beijing, China. Environ. Earth Sci. 69, 1571–1577.
Yuan, D.G., Zhang, G.L., Gong, Z.T., Burghardt, W., 2007. Variations of soil phosphorus accumulation in Nanjing, China as affected by urban development. J. Plant Nutr. Soil Sci. 170, 244–249.
Zhang, C., 2006. Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway. Irel. Environ. Pollut. 142, 501–511.
Zhang, C., Luo, L., Xu, W., Ledwith, V., 2008. Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway. Irel. Sci. Total Environ. 398, 212–221.
Zhang, C., Manheim, F.T., Hinde, J., Grossman, J.N., 2005. Statistical characterization of a large geochemical database and effect of sample size. Appl. Geochem 20, 1857–1874.
Zhang, C., Tang, Y., Xu, X., Kiely, G., 2011. Towards spatial geochemical modelling: use of geographically weighted regression for mapping soil organic carbon contents in Ireland. Appl. Geochem 26, 1239–1248.
Zhang, G., Burghardt, W., Lu, Y., Gong, Z.-T., 2001. Phosphorus-enriched soils of urban and suburban Nanjing and their effect on the groundwater phosphorus. J. Plant Nutr. Soil Sci. 164, 295–301.
Zhang, M.K., 2004. Phosphorus accumulation in soils along an urban-rural land use gradient in Hangzhou, Southeast China. Commun. Soil Sci. Plant Anal. 35, 819–833.

4.2 Submillimeter-scale heterogeneity of labile P in sediments characterized by the DGT and spatial analysis

Meng, Y., Ding, S., Gong, M., Chen, M., Wang, Y., Fan, X., Shi, L., Zhang, C., 2018. Submillimeter-scale heterogeneity of labile phosphorus in sediments characterized by diffusive gradients in thin films and spatial analysis. Chemosphere 194, 614-621.

Summary:

Sediments have a heterogeneous distribution of labile redox-sensitive elements which have been recognized to affect the labile P content. The DGT technique enables one to measure labile P in situ on a 2D submillimeter level. Conventionally, the 2D profile is displayed into 1D graph whose X axis is P concentration and Y axis is depth. This kind of statistic description would lose a lot of information, especially the horizonal spatial variation. This paper employed GIS techniques to visualize the labile P distribution at such a micro-scale, showing that the labile P was low in winter and high in summer. Spatial analysis methods, including semivariogram and Moran's I, were used to quantify the spatial variation of labile P. The Moran's I values were near one, indicating that the distribution had significant spatial autocorrelation during the whole year. The semivariograms illustrated clear submillimeter-scale spatial patterns with seasonal changes. The strong spatial variation in summer and weak spatial variation in winter suggested the influence of temperature on the mobility and spatial distribution of labile P in sediments.

My dedication in this paper accounted for 80% in reviewing literatures, exploring data and writing manuscript.

Chemosphere 194 (2018) 614–621

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Submillimeter-scale heterogeneity of labile phosphorus in sediments characterized by diffusive gradients in thin films and spatial analysis

Yuting Meng [a], Shiming Ding [b, *], Mengdan Gong [b], Musong Chen [b], Yan Wang [b, c], Xianfang Fan [b], Lei Shi [b], Chaosheng Zhang [a, **]

[a] International Network for Environment and Health, School of Geography and Archaeology, National University of Ireland, Galway, Ireland
[b] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China
[c] University of Chinese Academy of Sciences, Beijing 100049, China

HIGHLIGHTS

- Spatial distribution of labile P was visualized using GIS techniques.
- Spatial variation of labile P was quantified using semivariogram and Moran's I.
- Labile P had significant submillimeter-scale spatial autocorrelation.
- High values of labile P with strong spatial variation were observed in summer.

ARTICLE INFO

Article history:
Available online 5 December 2017

Handling Editor: X. Cao

Keywords:
Geographical information system (GIS)
Semivariogram
Moran's I
Lake Taihu

ABSTRACT

Sediments have a heterogeneous distribution of labile redox-sensitive elements due to a drastic downward transition from oxic to anoxic condition as a result of organic matter degradation. Characterization of the heterogeneous nature of sediments is vital for understanding of small-scale biogeochemical processes. However, there are limited reports on the related specialized methodology. In this study, the monthly distributions of labile phosphorus (P), a redox-sensitive limiting nutrient, were measured in the eutrophic Lake Taihu by Zr-oxide diffusive gradients in thin films (Zr-oxide DGT) on a two-dimensional (2D) submillimeter level. Geographical information system (GIS) techniques were used to visualize the labile P distribution at such a micro-scale, showing that the DGT-labile P was low in winter and high in summer. Spatial analysis methods, including semivariogram and Moran's I, were used to quantify the spatial variation of DGT-labile P. The distribution of DGT-labile P had clear submillimeter-scale spatial patterns with significant spatial autocorrelation during the whole year and displayed seasonal changes. High values of labile P with strong spatial variation were observed in summer, while low values of labile P with relatively uniform spatial patterns were detected in winter, demonstrating the strong influences of temperature on the mobility and spatial distribution of P in sediment profiles.

1. Introduction

Sediments are often considered as one-dimensional (1D) systems in which the redox processes are laterally uniform and sequentially layered. However, the development and application of high-resolution sampling techniques, such as diffusive gradients in thin films (DGT) and planar optodes, have revealed distinct heterogeneity in distributions of mobile chemical species (solutes in pore water and/or labile fractions associated with sediment solids) on the two-dimensional (2D), submillimeter level (Santner et al., 2015). Studies on the biochemistry of the sediments at a 2D, small scale have significantly promoted our understanding of the early diagenetic processes as well as the fate of anthropogenic contaminants (Stockdale et al., 2009).

Phosphorus (P) is an essential nutrient element for biotics and has long been recognized as a key limiting nutrient in freshwater

* Corresponding author..
** Corresponding author..
E-mail addresses: smding@niglas.ac.cn (S. Ding), chaosheng.zhang@nuigalway.ie (C. Zhang).

https://doi.org/10.1016/j.chemosphere.2017.11.178
0045-6535/

systems (Schindler, 1977; Schindler et al., 2016). It is known that the release of P from sediments, i.e., the internal loading of P, can contribute considerably to the water column P and can be a major driving force for water eutrophication (Søndergaard et al., 2013; Nürnberg and LaZerte, 2016; Wu et al., 2017). The release of P is primarily controlled by the iron (Fe) redox cycle (Ding et al., 2016). There is an oxic layer in the uppermost sediment layer composed of mainly Fe oxyhydroxides, which retain P and inhibit its release to the water column. In deeper sediment layers under anoxic condition, reductive dissolution of Fe oxyhydroxides occurs, resulting in the simultaneous release of Fe(II) and P. The released Fe(II) and P will diffuse upward and be immobilized in the oxic layer at the sediment-water interface (SWI). Besides the Fe redox cycle, many other processes affect the mobility of P in sediments, including aluminum hydroxide adsorption (Lake et al., 2007; Lin et al., 2017), calcium-bound P dissolution (Golterman, 2001), organic matter degradation (Lü et al., 2016), and bacterial P release (Maitra et al., 2015). For example, it has been found that the degradation of algae settled on the SWI can directly release P to the water column and decouple the relationship between Fe redox and P release processes (Han et al., 2015). As the redox conditions change sharply from oxic to anoxic in heterogeneous sediments, these processes have a combined effect on the mobility of P in sediments and may cause significant variation in the distribution of labile P (Zou et al., 2017). The heterogeneous distribution of labile P in sediments has been revealed on a 2D, submillimeter level, using DGT measurement (Stockdale et al., 2008; Ding et al., 2011, 2013). The horizontal heterogeneity of labile P in the eutrophic Lake Taihu has been demonstrated by the ratio of standard deviation to mean concentration of labile P at each depth, and the value varied from 0.04 to 4.5. High values appeared at the depth of 30 mm below the SWI, demonstrating that the remobilization of P is active at these depths (Ding et al., 2015).

The 2D-DGT technique makes it possible to distinguish between two types of spatial distributions of labile P: (i) random distributions that are associated with benthic fauna or bioturbation (Adámek and Maršálek, 2013); (ii) spatial autocorrelations that are related to environmental factors, such as dissolved oxygen (DO), pH, temperature and other relevant chemical parameters (Christophoridis and Fytianos, 2006). The spatial autocorrelation means that the properties for locations close together are similar, while for large distances they are statistically independent. In this case, the classic statistical methods for spatially independent data, such as calculation of arithmetic mean and standard deviation, are no longer sufficient (Niemietz et al., 2010). Instead, the spatial analysis could be applied in such analysis which takes the spatial autocorrelations into consideration.

The semivariogram and kriging method, initially used for mining industry, is now widely employed in environment sciences (Olive and Webster, 2015). It is one of the most frequently used statistical methods to evaluate the spatial autocorrelation. Semivariogram modeling can quantify the degree of dissimilarity of a pair of values at different locations, which makes it the basis of many geostatistical applications, such as optimum spatial sampling strategy and interpolation (Liu et al., 2013). Besides, spatial autocorrelation can also be quantified by Moran's I, which represents the deviation among the values of the variable and its mean (Moran, 1950; Kim et al., 2003).

In the present study, the monthly distribution of labile P over the period of February 2016 to January 2017 were measured in the eutrophic Lake Taihu by Zr-oxide DGT on a 2D, submillimeter level. The main objectives of this study were (i) to visualize the monthly distribution of labile P in sediments using geographical information system (GIS) techniques so as to compare the difference of distribution through time series, (ii) to quantify spatial autocorrelation of labile P using spatial analysis methods, including semivariogram and Moran's I, in order to determine the spatial structure characteristics of labile P in sediments of Lake Taihu.

2. Materials and methods

2.1. Study area

Lake Taihu (31°18′42.7″ N, 119°56′52.2″ E) (Fig. 1) is located in south-eastern China with an annual average water surface area of 2338 km^2. As the third largest freshwater lake in China, Lake Taihu is a water source for several large cities including Suzhou, Wuxi and Huzhou. However, cyanobacterial blooms have occurred frequently over recent years due to intensive industry and agriculture development which has brought excessive nutrients to Lake Taihu in the past decades (Xu et al., 2013). The sampling site is situated in Meiliang Bay, northwest region of Lake Taihu, which is severely affected by cyanobacterial blooms (Li et al., 2016).

2.2. DGT theory

The fundamental theory behind DGT is Fick's first law of diffusion (Davison and Zhang, 2012). The labile forms of chemical elements diffuse through the diffusion layer (composed of a filter membrane and a diffusive gel) and are then bound on the binding gel. The concentration of P can be calculated using Equation (1) (Davison and Zhang, 2012):

$$C_{DGT} = \frac{M\Delta g}{DAt} \quad (1)$$

where Δg (cm, typically 0.9 mm) represents the thickness of the diffusive layer, D ($cm^2\ s^{-1}$) refers to the diffusive coefficient of P in the diffusive layer, t (s) is the deployment time, A (cm^2) represents the exposure area of the gel, and M (μg) is the accumulated mass over the deployment time (e.g., 24 h).

For the purpose of submillimeter measurement, a thin filter membrane is generally used as the diffusion layer. In this case, the measured C_{DGT} was far lower than the pore water concentration, and the DGT-labile P was represented as flux (pg $cm^{-2}\ s^{-1}$) (Widerlund et al., 2012). The Zr-oxide DGT probes used in this study were made of high-capacity Zr-oxide binding gel, which was prepared according to the procedure described by Ding et al. (2011). A PVDF membrane with a pore size of 0.45 μm and a thickness of 0.10 mm (Millipore, USA) was used as the diffusion layer. The Zr-oxide DGT probes were deoxygenated with nitrogen for 16 h before deployment.

2.3. Sampling

Three parallel sediment cores were collected into plexiglass tubes using a gravity core sampler once a month from February 2016 to January 2017 at sampling site of Meiliang Bay, Lake Taihu. The core sediments were transported to laboratory and deployed overnight, while the temperature was kept the same as in the field. Then, a Zr-oxide DGT probe was inserted into one core and deployed for 24 h. After retrieval, the SWI was marked immediately, and then each probe was rinsed using deionized water. The accumulated mass of P in the binding gel was determined according to the method reported by Ding et al. (2013). Each binding gel was heated in hot water (85 °C) for 5 d, and was then immersed in a vessel with the mixed reagent (ammonium molybdate tetrahydrate and potassium antimonyl tartrate), and maintained at 35 °C for 45 min. Finally, the surface of the gel was scanned by a flat-bed scanner at a resolution of 600 dpi, corresponding to a pixel size

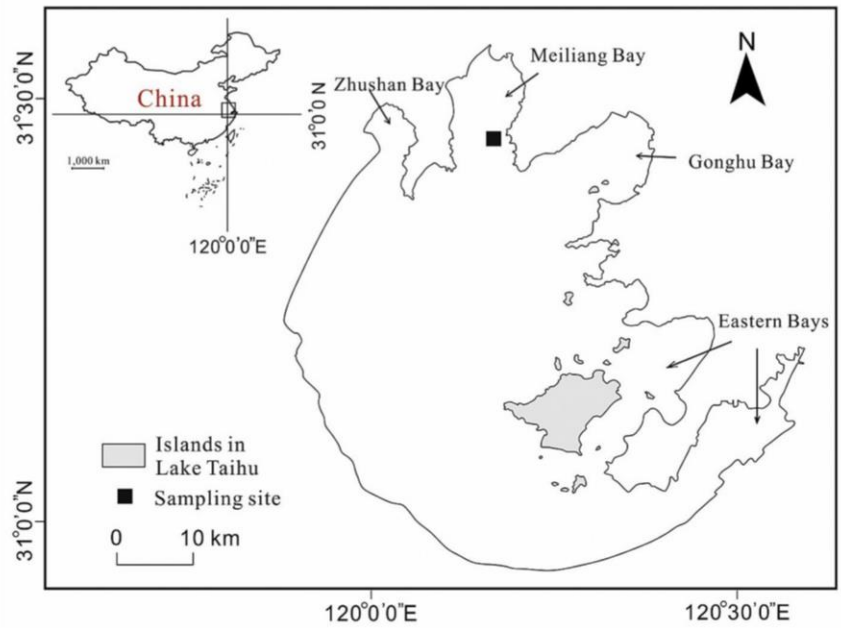


Fig. 1. Sampling site in Lake Taihu.

of 42 × 42 μm. The grayscale intensity was further transferred into flux (accumulated mass per area) according to Equation (2) (Ding et al., 2013):

$$G = -167.29e^{-\frac{M}{6.51}} + 214.63 \quad (2)$$

where G is the grayscale intensity of the gel surface, and M (μg) is the accumulated mass over the deployment time (e.g., 24 h).

2.4. Semivariogram

Semivariogram analysis was performed for DGT-labile P in sediments according to Equation (3) (Cressie, 1993):

$$r(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (3)$$

where $r(h)$ is the semivariogram for a lag distance h between a pair of observations $Z(x_i)$ and $Z(x_i + h)$, and $N(h)$ is the number of data pairs separated by lag distance h. The results in this study were best fitted by the Gaussian model (Isaaks and Srivastava, 1989):

The vertical intercept is called nugget, which can be attributed to errors of measurement and spatial microvariability at distances smaller than the sampling interval (Niemietz et al., 2010). Semivariance increases smoothly as the separation distance increases, indicating that samples which are located closer to each other exhibit less difference values of the measured property. The sill, which is the upper bound of variance, is constant for some semivariograms and an asymptote for others. The sill means the variation level. The range is the distance at which the sill reaches stable constant or asymptotic, marking the maximum spatial autocorrelation distance. Locations further apart than this distance are spatially uncorrelated or independent. If the sill is asymptotic, an effective range where the variance reaches 95% of its sill can be used (Olive and Webster, 2015).

2.5. Moran's I

Moran's I is calculated by Equation (4) (Cliff and Ord, 1981):

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \times \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (4)$$

where n is the total number of the variable, x_i and x_j are the values of the variable x at the location i and j, $\bar{x}$ is the average and w_{ij} is the weight function. The weight function defines a local neighborhood around the location i. The value at the location i is compared with the weighted average of the values of its neighbors. Usually, Moran's I value ranges from −1 to 1. The value between 0 and 1 represents positive autocorrelation (with same high value or low value patterns); the value between −1 and 0 means negative autocorrelation (with different high and low value patterns in the distance band); and a value close to zero represents spatial randomness. To test the significance level of spatial autocorrelation, the Z score for Moran's I is often calculated and standardized. In this study, data sets were determined by a threshold of 1.96 with a 95% confidence level and the results were based on 9999 permutations to avoid sensitivity on the specific randomization (Huo et al., 2012). A Z score >1.96 presents significant positive autocorrelation indicating that the spatial distribution of the variable under study contains clear patterns with clusters of high values or lower values, instead of being randomly distributed.

Correlogram is a graph of Moran's I values plotted against the varying distances between observations which is often used for checking randomness in a data set. If the autocorrelation fluctuates near zero for all distance lag, the data is random. If the autocorrelation shows a significant departure from zero, the data has spatial autocorrelation.

2.6. Data analyses and computer software

Each scanned grayscale image comprised 472 × 2829 pixels (for

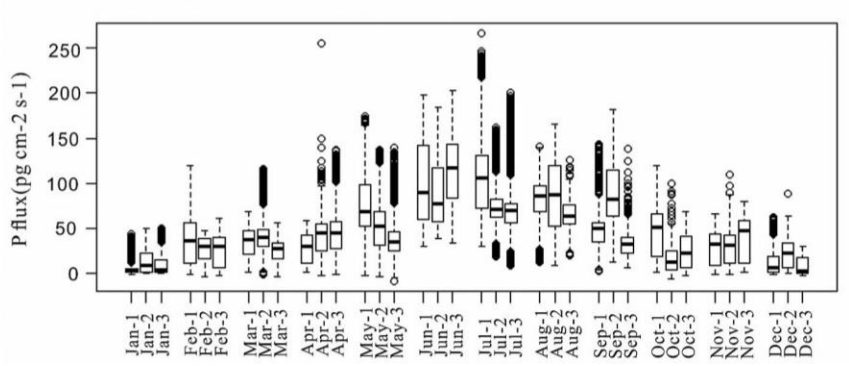


Fig. 2. Boxplot diagram of labile P flux in sediments of Meiliang Bay, Lake Taihu.

most profiles in this study). The grayscale images (raster data) were then changed to point data in order to obtain the values of greyscale intensity using ArcGIS (version 10.3). The values obtained from each grayscale image contained 472 columns and 2829 rows, 1,335,288 points in total, which were calculated to P flux using Equation (2) in ArcGIS. The basic statistics were computed using the 472 × 2829 values of P flux. The distribution maps were produced using Point to Raster Tool through inputting point data shapefile and filling P flux into value field in ArcGIS. The semivariograms and Moran's I analysis were conducted based on the aggregated data because the point data were too large for calculation. Recently, a new problem has emerged that computer cannot handle data larger than its memory (Yu et al., 2012) at the "big data" era. The aggregated data were produced using raster package in free R program (version 3.4.1), with the aggregation of 10 × 10 points into 1 point based on the mean value. As a result, the size of each aggregated data set was 48 columns and 283 rows with a resolution of 0.42 × 0.42 mm. All maps were produced using ArcGIS software.

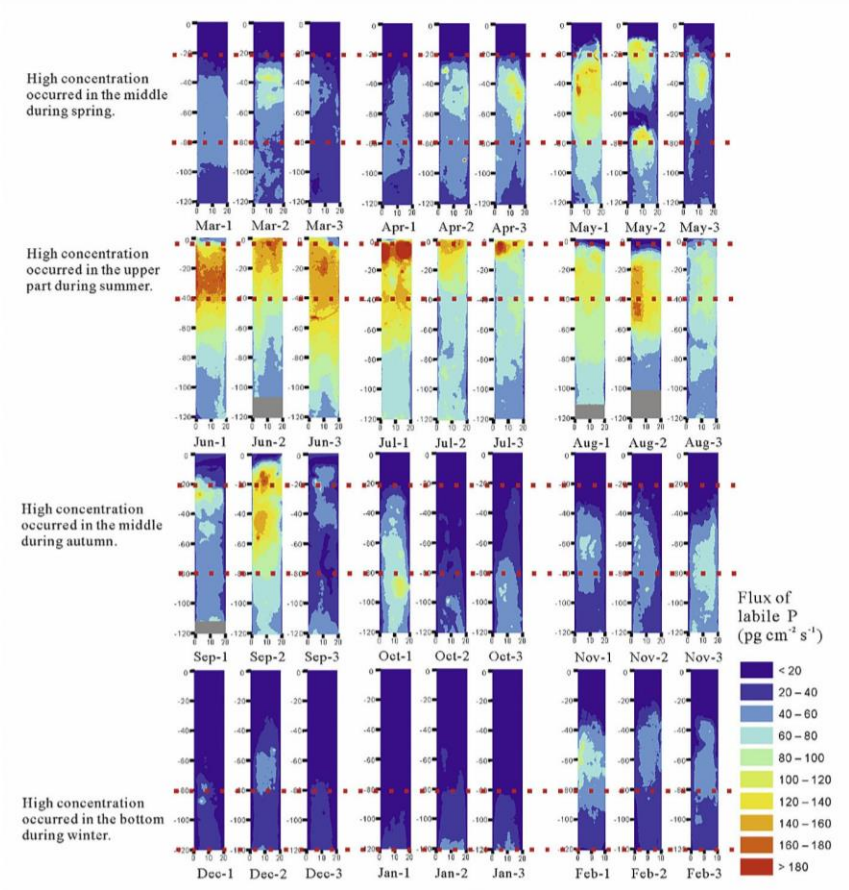


Fig. 3. 2D distribution of DGT P flux in sediments of Meiliang Bay, Lake Taihu.

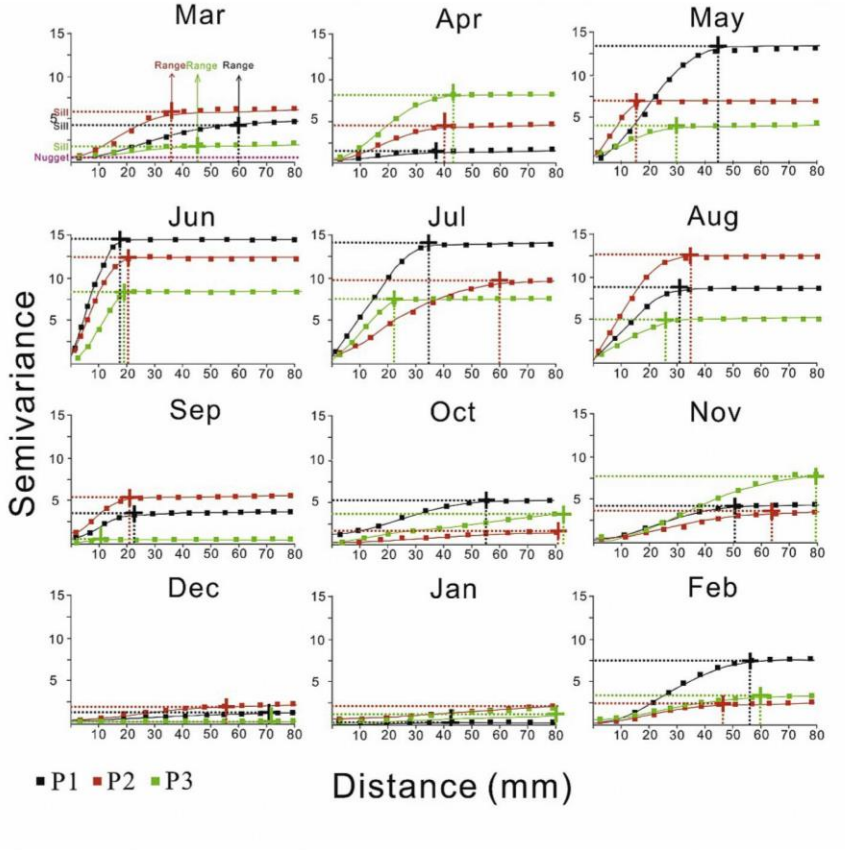


Fig. 4. Semivariograms for DGT-labile P in sediments of Meiliang Bay, Lake Taihu. (The labels P1, P2, and P3 refer to parallel samplings in each month).

The semivariograms were obtained in ArcGIS. The Moran's I was computed using spatial autocorrelation tool of spatial statistics tools in ArcGIS. The conceptualization of spatial relationship of this tool was chosen as "Inverse Distance". Meaning that nearby neighboring values have a larger weight on the computations for a target value than values that are far away. Selecting a distance band value as a moving window is to some extent arbitrary. In this case, 10 mm (the minimum of the ranges) was selected as the moving window. The spatial correlograms were calculated using free R program.

3. Results and discussion

3.1. Basic statistics

The skewness values, which can be roughly observed in the boxplots shown in Fig. 2, were mostly close to 0, with slightly higher values for January-1&3, July-2&3 and December-1 (The numbers show the parallel samples in each month). Therefore, the raw aggregated data were used in order to compare semivariograms and Moran's I indices among all the 36 data sets. Lalibe P was high in summer, low in winter and medium in spring and autumn. The smallest mean labile P was observed in January-1 data set (6.40 pg cm^{-2} s^{-1}) and the largest mean labile P ocurred in June-3 data set (112.52 pg cm^{-2} s^{-1}) (Fig. 2).

3.2. Spatial distribution of labile P

The 2D distribution of DGT-labile P in sediments of Meiliang Bay, Lake Taihu was produced by ArcGIS (Fig. 3). In general, with the increase in depth, labile P increased gradually at first, and then stabilized or decreased later. Some edges of maps were biased by the "edge effect" of the edge of the DGT device itself. In summer, labile P and its spatial variation were higher than that in other seasons. When the temperature was high, the activities of microorganism and benthic fauna were enhanced, resulting in rapid decomposition of organic matter. With the depletion of the oxygen, Fe(II) which was reduced by Fe(III) promoted the release of P (Andersen and Ring, 1999). The peak of labile P existed in the upper part of the sediments in summer. However, the most top (0–5 mm) of sediments had low labile P, related to the edge effect and the oxic condition on the top near SWI. In spring and autumn, the peak of labile P mainly occurred in the middle of the map, corresponding to the thickening of oxic layer where abundant Fe oxide absorbed P strongly (Li et al., 2010). During the winter months, labile P was much lower and the depth of maximum labile P moved down further due to strengthened oxic conditions (Ding et al., 2016). The spatial variation was very small and labile P was relatively uniform in winter.

3.3. Semivariance analysis of DGT-labile P

A semivariogram for each data set was calculated to describe the spatial variation of labile P (Fig. 4). The semivariogram model clearly illustrated the nugget, sill and range of each data set (Table 1), indicating that the labile P distributions had significant spatial autocorrelation rather than randomness. The results of semivariogram were consistent with the 2D distribution of labile P

Table 1
Spatial properties of the semivariogram parameters, Moran's I and standardized Z values of DGT-labile P in sediments of Meiliang Bay, Lake Taihu.

Time	Nugget	Sill	Range (mm)	N/S ratio (%)	Moran's I	Z score
January-1	0.07	0.34	41.91	20.59	0.870	768
January-2	0.08	2.14	88.65	3.74	0.908	802
January-3	0.07	1.5	80.01	4.67	0.922	814
February-1	0.09	7.51	57.15	1.20	0.931	776
February-2	0.11	2.23	47.24	4.93	0.893	744
February-3	0.45	3.07	59.18	14.66	0.837	697
March-1	0.4	3.69	61.21	10.84	0.953	842
March-2	0.23	5.27	37.34	4.36	0.856	756
March-3	0.1	1.71	45.21	5.85	0.872	771
April-1	0.29	1.75	38.86	16.57	0.810	716
April-2	0.3	4.24	40.64	7.08	0.811	717
April-3	0.57	7.91	42.16	7.21	0.834	737
May-1	0.72	13.8	43.43	5.23	0.864	763
May-2	0.47	6.89	15.24	6.82	0.712	629
May-3	0.76	3.94	29.21	19.29	0.731	646
June-1	0.57	14.8	17.91	3.85	0.929	931
June-2	0.49	12.24	20.70	4.00	0.964	966
June-3	0.65	8.34	19.69	7.80	0.916	918
July-1	1.66	13.9	33.31	11.94	0.880	872
July-2	0.84	9.78	59.44	8.59	0.821	814
July-3	0.51	7.56	22.86	6.75	0.789	783
August-1	0.53	8.85	31.24	5.99	0.763	815
August-2	0.33	12.66	34.88	2.60	0.891	864
August-3	0.42	5.07	25.91	8.28	0.674	669
September-1	0.2	3.51	22.86	5.70	0.787	467
September-2	1.16	5.86	20.83	19.80	0.798	505
September-3	0.09	0.76	10.41	11.84	0.782	481
October-1	1.19	5.42	55.37	21.96	0.810	813
October-2	0.19	1.9	81.28	10.00	0.878	515
October-3	0.54	3.74	83.31	14.44	0.857	726
November-1	0.24	4.04	50.55	5.94	0.882	521
November-2	0.45	3.23	63.75	13.93	0.897	596
November-3	0.22	7.71	79.50	2.85	0.935	540
December-1	0.3	1.14	71.12	26.32	0.874	501
December-2	0.31	2.02	56.90	15.35	0.867	497
December-3	0.15	1	71.88	15.00	0.854	584

The labels 1, 2, and 3 refer to parallel samplings in each month. N/S: Nugget/Sill

in sediments. In this study, the nugget values for all semivariograms were small, indicating that the measurement error and the spatial variability at smaller scales were mall, which were attributed to the precise measurement and high resolution of the data sets. Relatively high sill values were observed in summer, implying large variation of labile P in this season, while the sill values in winter were low with small variation of labile P results. Meanwhile, none of the semivariograms exhibited a parabolic feature of non-stationary, as all of them reached a relative flat shape (plateau) in the long intervals. Another important parameters in the semivariogram model is range, as it indicates the samples' spatial correlation at specific intervals. This distance could be interpreted as average patch size. The term "patch" presents an area that differs from its surroundings, but is not necessarily internally homogeneous (Dent and Grimm, 1999). For example, the spatial autocorrelation of June and September disappeared when the distance was more than approximately 20 mm, whereas the spatial autocorrelation of January and December vanished when the interval was above ~70 mm. The nugget/sill ratio roughly defines different classes of spatial correlation (Fu et al., 2011). The results show that a strong spatial correlation (nugget/sill ratio < 25%) existed in all data sets, except for August-3 and December-1. The strength of spatial dependence can be measured by the slope of the ascending curve of the semivariogram. The slopes of winter were flat, indicating strong spatial dependences, while the slopes of summer were steep, suggesting weak spatial dependences.

In summary, the semivariograms varied monthly and showed a seasonal pattern. The ranges of winter were large and their slopes were flat, indicating that the distribution of DGT-labile P had big patches and were nearly uniform. Spring and autumn shared a similar pattern (moderate spatial correlation ranges and moderate slopes), suggesting more patchy distribution and moderate spatial variation. Summer was characterized by small patches size and high heterogeneity which was indicated by short ranges and steep slopes. Besides, strong spatial variations were also observed in May and September, except for summer, indicating that seasonal change had strong influence on the distribution of labile P. Widerlund and Davison (2007) also reported seasonal increase of microniches measured by DGT. The intensive variability of labile P distribution in May was in agreement with a short-term, high-magnitude release of sediment P to the overlying water after the onset of anoxia observed by Ding et al. (2016). The strong spatial variations occurring in the spring and autumn may be a result of the rapid change of Fe oxyhydroxides triggered by the fluctuant oxic and anoxic conditions at SWI.

3.4. Moran's I analysis of DGT-labile P

Table 1 also shows the Moran's I values and standardized Z scores of each data set which was weighted by inverse distance with a distance band of 10 mm. During optimization for the weight method, we found that when threshold distance was small, the value of Moran's I was very close to 1 and there was no obvious difference among the data set. Based on the semivariograms, the minimum of range was about 10 mm. Therefore, threshold distance was defined as 10 mm in order to observe differences among the data sets more clearly. All data sets had good spatial patterns because of significantly positive autocorrelation ($p \leq 0.05$) suggested by the large Z scores (far above 1.96). To further understand the spatial autocorrelation, a spatial correlogram was produced for each data set, which is a graph of Moran's I values plotted against the distances between observations (Fig. 5). It was seen that the correlograms were in accordance for the three parallel data sets of each month. Positive autocorrelation (high-high and low-low distribution), representing as the first x-intercept, reflects the average radius of spatial patch, whereas negative autocorrelation (high-low distribution) reflects the average distance between the opposite attribute spatial patches (Zhang and Selinus, 1997). In general, spatial correlation weakened as distance lag increased. Summer had small size of patches; spring and autumn had middle size of patches; winter had large size of patches (Fig. 6). The size of patch indicates the degree of spatial variation. The smaller the patch is, the more heterogeneous the spatial pattern is, vice versa. The correlogram results were similar to the semivariogram results, however, the correlogram can distinguish between positive and negative autocorrelation. Considering the correlogram, the positive autocorrelation distance was about 0–40 mm and the negative autocorrelation distance was about 40–60 mm in winter; the positive autocorrelation distance was about 0–30 mm and the negative autocorrelation distance was about 30–60 mm in summer; the positive autocorrelation distance was about 0–35 mm and the negative autocorrelation distance was about 35–60 mm in spring and autumn. Interestingly, May and September possessed small Moran's I and small spatial patch of nearly 20 mm. The correlogram of September even fluctuated around zero when the distance lag was larger than 20 mm, indicating relatively randomness of spatial clusters and spatial outliers over the entire region. This relatively random spatial structure was likely affected by random process, such as benthic fauna or bioturbation (Chen et al., 2016a, 2016b).

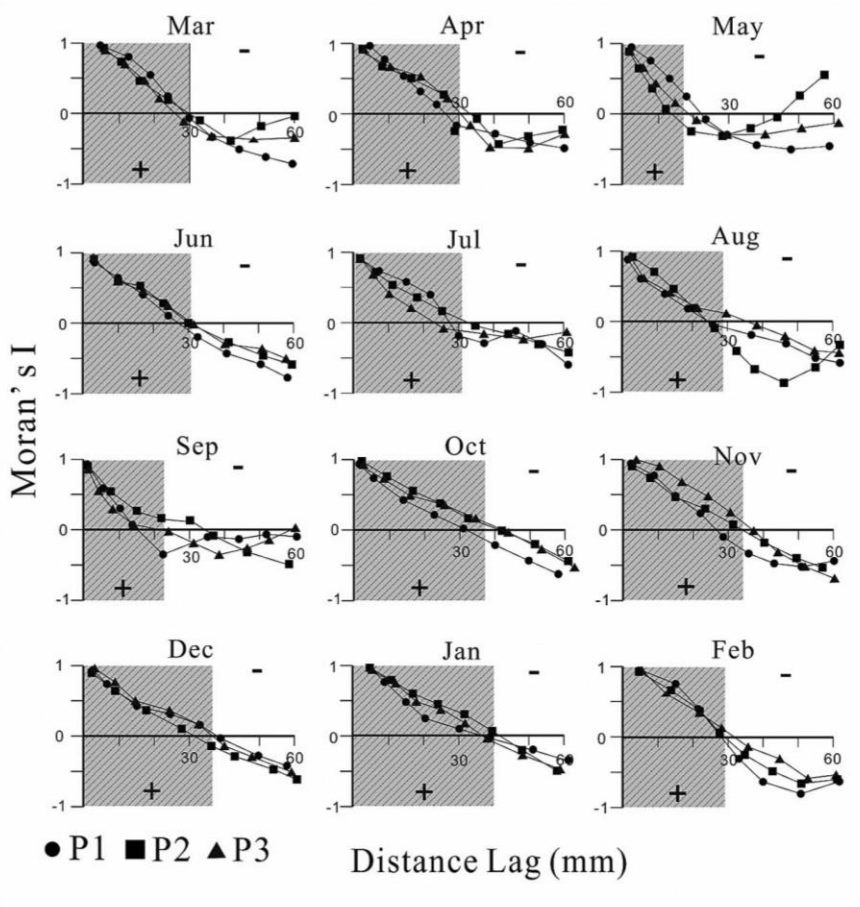


Fig. 5. Spatial correlograms of DGT-labile P in sediments of Meiliang Bay, Lake Taihu. (The labels P1, P2, and P3 refer to parallel samplings in each month).

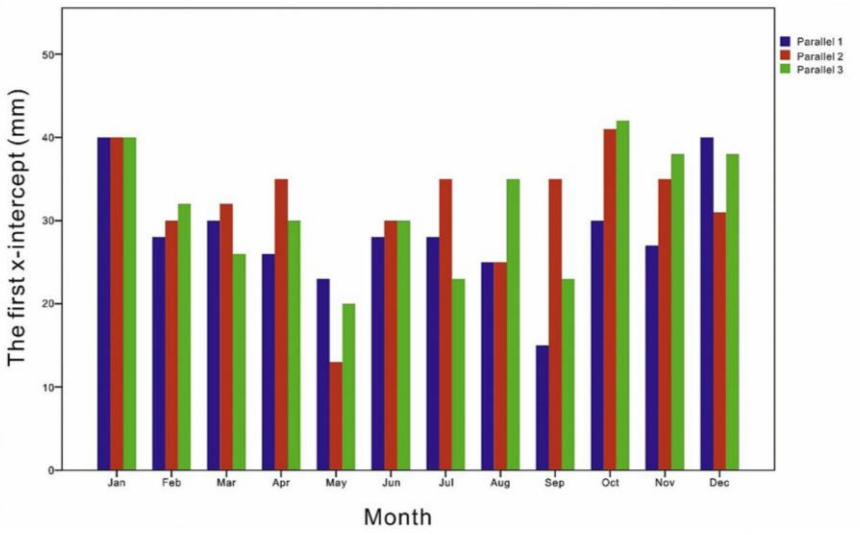


Fig. 6. Bar chart of the first x-intercepts of correlograms.

4. Conclusions

The maps produced using GIS techniques showed remarkable difference of mass and distribution of labile P in sediment sections. Winter had low labile P; spring and autumn had moderate labile P; and summer had high labile P. The maximum labile P of winter was

at the bottom of the map; the maximum labile P of spring and autumn was at the middle of the map; and the maximum labile P of summer was at the top of the map. This paper reveals that spatial analysis, based on semivariogram and Moran's I, provides useful information about the 2D, submillimeter-scale spatial variation of labile P in sediments. It is shown that there exists a strong spatial autocorrelation of labile P, in spite of seasonal change of spatial autocorrelation. In winter, the spatial autocorrelation of labile P with low spatial variations had large patches, suggesting strong spatial dependence and relatively uniform spatial pattern. In summer, the spatial autocorrelation of labile P with strong spatial variations had small patches, indicating weak spatial dependence and complex spatial pattern.

Acknowledgements

This study was jointly sponsored by the National Scientific Foundation of China (41621002, 41571465, 41322011) and National High-level Personnel of Special Support Program.

References

Adámek, Z., Maršálek, B., 2013. Bioturbation of sediments by benthic macroinvertebrates and fish and its implication for pond ecosystems: a review. Aquac. Int. 21, 1–17.
Andersen, F.Ø., Ring, P., 1999. Comparison of phosphorus release from littoral and profundal sediments in a shallow, eutrophic lake. Hydrobiologia 408–409, 175–183.
Chen, M., Ding, S., Liu, L., Wang, Y., Xing, X., Wang, D., Gong, M., Zhang, C., 2016a. Fine-scale bioturbation effects of tubificid worm (Limnodrilus hoffmeisteri) on the lability of phosphorus in sediments. Environ. Pollut. 219, 604–611.
Chen, M., Ding, S., Liu, L., Xu, D., Gong, M., Tang, H., Zhang, C., 2016b. Kinetics of phosphorus release from sediments and its relationship with iron speciation influenced by the mussel (Corbicula fluminea) bioturbation. Sci. Total Environ. 542 (Part A), 833–840.
Christophoridis, C., Fytianos, K., 2006. Conditions affecting the release of phosphorus from surface lake sediments. J. Environ. Qual. 35, 1181–1192.
Cliff, A.D., Ord, J.K., 1981. Spatial Processes: Models and Applications. Pion Limited, London.
Cressie, N.A.C., 1993. Statistics for Spatial Data. Revised Edition Edn.. Wiley, New York.
Davison, W., Zhang, H., 2012. Progress in understanding the use of diffusive gradients in thin films (DGT) – back to basics. Environ. Chem. 9, 1–13.
Dent, C.L., Grimm, N.B., 1999. Spatial heterogeneity of stream water nutrient concentrations over successional time. Ecology 80 (7), 2283–2298.
Ding, S., Han, C., Wang, Y., Yao, L., Wang, Y., Xu, D., Sun, Q., Williams, P.N., Zhang, C., 2015. In situ, high-resolution imaging of labile phosphorus in sediments of a large eutrophic lake. Water Res. 74, 100–109.
Ding, S., Jia, F., Xu, D., Sun, Q., Zhang, L., Fan, C., Zhang, C., 2011. High-resolution, two-dimensional measurement of dissolved reactive phosphorus in sediments using the diffusive gradients in thin films technique in combination with a routine procedure. Environ. Sci. Technol. 45 (22), 9680–9686.
Ding, S., Wang, Y., Wang, D., Li, Y., Gong, M., Zhang, C., 2016. In situ, high-resolution evidence for iron-coupled mobilization of phosphorus sediment. Sci. Rep. 6, 24341.
Ding, S., Wang, Y., Xu, D., Zhu, C., Zhang, C., 2013. Gel-based coloration technique for the submillimeter-scale imaging of labile phosphorus in sediments and soils with diffusive gradients in thin films. Environ. Sci. Technol. 47 (14), 7821–7829.
Fu, W., Zhao, K., Zhang, C., Tunney, H., 2011. Using Moran's I and geostatistics to identify spatial patterns of soil nutrients in two different long-term phosphorus-application plots. J. Plant Nutr. Soil Sci. 174, 785–798.
Golterman, H.L., 2001. Phosphate release from anoxic sediments or 'What did Mortimer really write?'. Hydrobiologia 450, 99–106.
Han, C., Ding, S., Yao, L., Shen, Q., Zhu, C., Wang, Y., Xu, D., 2015. Dynamics of phosphorus–iron–sulfur at the sediment–water interface influenced by algae blooms decomposition. J. Hazard. Mater. 300, 329–337.
Huo, X.N., Li, H., Sun, D.F., Zhou, L.D., Li, B.G., 2012. Combining geostatistics with Moran's I analysis for mapping soil heavy metals in Beijing, China. Int. J. Environ. Res. Public Health 9 (3), 995–1017.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press.
Kim, L.H., Choi, E., Stenstrom, M.K., 2003. Sediment characteristics, phosphorus types and phosphorus release rates between river and lake sediments. Chemosphere 50 (1), 53–61.
Lü, C., He, J., Zuo, L., Vogt, R.D., Zhu, L., Zhou, B., Mohr, C.W., Guan, R., Wang, W., Yan, D., 2016. Processes and their explanatory factors governing distribution of organic phosphorous pools in lake sediments. Chemosphere 145, 125–134.
Lake, B.A., Coolidge, K.M., Norton, S.A., Amirbahman, A., 2007. Factors contributing to the internal loading of phosphorus from anoxic sediments in six Maine, USA, lakes. Sci. Total Environ. 373, 534–541.
Li, D., Erickson, R.A., Tang, S., Zhang, Y., Niu, Z., Liu, H., Yu, H., 2016. Structure and spatial patterns of macrobenthic community in Taihu Lake, a large Shallow lake, China. Ecol. Indic. 61 (part 2), 179–187.
Li, Y., Xia, B., Zhang, J., Li, C., Zhu, W., 2010. Assessing high resolution oxidation-reduction potential and soluble reactive phosphorus variation across vertical sediments and water layers in Xinghu Lake: a novel laboratory approach. J. Environ. Sci. 22 (7), 982–990.
Lin, J., Sun, Q., Ding, S., Wang, D., Wang, Y., Chen, M., Shi, L., Fan, X., Tsang, D.C.W., 2017. Mobile phosphorus stratification in sediments by aluminum immobilization. Chemosphere 186, 644–651.
Liu, Q., Xie, W., Xia, J., 2013. Using semivariogram and Moran's I techniques to evaluate spatial distribution of soil micronutrients. Commun. Soil Sci. Plant Anal. 44, 1182–1192.
Maitra, N., Manna, S.K., Samanta, S., Sarkar, K., Debnath, D., Bandopadhyay, C., Sahu, S.K., Sharma, A.P., 2015. Ecological significance and phosphorus release potential of phosphate solubilizing bacteria in freshwater ecosystems. Hydrobiologia 745, 69–83.
Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika 37, 17–23.
Nürnberg, G.K., LaZerte, B.D., 2016. More than 20 years of estimated internal phosphorus loading in polymictic, eutrophic Lake Winnipeg, Manitoba. J. Great Lakes Res. 42, 18–27.
Niemietz, K., Wagner, A., Gründig-Wendrock, B., Stoyan, D., Niklas, J.R., 2010. Variogram analysis of charge-carrier effective lifetime topograms in mc-Si materials. Sol. Energy Mater. Sol. Cells 94, 164–170.
Olive, M.A., Webster, R., 2015. Basic Steps in Geostatistics: the Variogram and Kriging. ISSN: 2211-808X. Springer, p. 1. https://doi.org/10.1007/978-3-319-15865-5.
Søndergaard, M., Bjerring, R., Jeppesen, E., 2013. Persistent internal phosphorus loading during summer in shallow eutrophic lakes. Hydrobiologia 710, 95–107.
Santner, J., Larsen, M., Kreuzeder, A., Glud, R.N., 2015. Two decades of chemical imaging of solutes in sediments and soils – a review. Anal. Chim. Acta 878, 9–42.
Schindler, D.W., 1977. Evolution of phosphorus limitation in lakes. Science 195, 260–262.
Schindler, D.W., Carpenter, S.R., Chapra, S.C., Hecky, R.E., Orihel, D.M., 2016. Reducing phosphorus to Curb Lake eutrophication is a success. Environ. Sci. Technol. 50, 8923–8929.
Stockdale, A., Davison, W., Zhang, H., 2008. (micro)High-resolution two-dimensional quantitative analysis of phosphorus, vanadium and arsenic, and qualitative analysis of sulfide, in a freshwater sediment. Envir. Chem. 5, 143–149.
Stockdale, A., Davison, W., Zhang, H., 2009. Micro-scale biogeochemical heterogeneity in sediments: a review of available technology and observed evidence. Earth Sci. Rev. 92, 81–97.
Wu, Z., Liu, Y., Liang, Z., Wu, S., Guo, H., 2017. Internal cycling, not external loading, decides the nutrient limitation in eutrophic lake: a dynamic model with temporal Bayesian hierarchical inference. Water Res. 116, 231–240.
Widerlund, A., Davison, W., 2007. Size and density distribution of sulfide-producing microniches in lake sediments. Environ. Sci. Technol. 41 (23), 8044–8049.
Widerlund, A., Nowell, G.M., Davison, W., Pearson, D.G., 2012. High-resolution measurements of sulphur isotope variations in sediment pore-waters by laser ablation multicollector inductively coupled plasma mass spectrometry. Chem. Geol. 291, 278–285.
Xu, D., Ding, S., Li, B., Bai, X., Fan, C., Zhang, C., 2013. Speciation of organic phosphorus in a sediment profile of Lake Taihu I: chemical forms and their transformation. J. Environ. Sci. 25 (4), 637–644.
Yu, H., Hsieh, C., Chang, K., Lin, C., 2012. Larger linear classification when data cannot fit in memory. ACM Trans. Knowl. Discov. Data 5 (4), 23 (February 2012), 23 pages.
Zhang, C., Selinus, O., 1997. Spatial analyses for copper, lead and zinc contents in sediments of the Yangtze River basin. Sci. Total Environ. 204, 251–262.
Zou, J., Zhang, L., Wang, L., Li, Y., 2017. Enhancing phosphorus release from waste activated sludge containing ferric or aluminum phosphates by EDTA addition during anaerobic fermentation process. Chemosphere 171, 601–608.

4.3 Identifying geogenic and anthropogenic controls on different spatial distribution patterns of Al, Ca and Pb in urban topsoil of the GLA area

Meng, Y., Cave, M., Zhang, C., 2020. Identifying geogenic and anthropogenic controls on different spatial distribution patterns of aluminium, calcium and lead in urban topsoil of Greater London Authority area. Chemosphere 238, 124541.

Summary:

Geochemical elements are influenced by human activities, especially in urban soils. However, some elements are resistant to variation and some are liable to variation. The significance of this study is that three chosen typical metal elements clearly demonstrate the different levels affected by human activities in an extremely urbanized city based on ANOVA and GIS-based spatial analysis. Understanding how much human activities influence the spatial variation is important for soil management. To conclude, Al is an inert element which can still maintain its natural spatial distribution even in the intensively urbanised GLA area; Pb is more easily affected by human activities which can be significantly changed; Ca demonstrates a mixed spatial distribution affected by both natural factors and human activities. In the geogenic aspect, Al tends to be enriched in clay PMs, and Ca is abundant in White Chalk Subgroup. In the anthropogenic aspect, building activity has a great impact on Ca distribution, while built-up and road traffic have strong influence on Pb distribution.

My dedication in this paper accounted for 90% in reviewing literatures, exploring data and writing manuscript.

Chemosphere 238 (2020) 124541

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Identifying geogenic and anthropogenic controls on different spatial distribution patterns of aluminium, calcium and lead in urban topsoil of Greater London Authority area

Yuting Meng [a], Mark Cave [b], Chaosheng Zhang [a, *]

[a] International Network for Environment and Health, School of Geography and Archaeology & Ryan Institute, National University of Ireland, Galway, Ireland
[b] British Geological Survey, Environmental Science Centre, Nottingham, United Kingdom

HIGHLIGHTS

- Spatial analyses were useful in identifying spatial patterns of urban geochemistry.
- Different levels of geogenic and anthropogenic controls found on different chemicals.
- Al was mainly controlled by geology, elevated in clay areas.
- Pb was mostly under anthropogenic control, elevated in built-up area and near road.
- Ca was impacted by both geogenic and anthropogenic controls.

ARTICLE INFO

Article history:
Received 30 May 2019
Received in revised form
5 August 2019
Accepted 7 August 2019
Available online 12 August 2019

Handling Editor: X. Cao

Keywords:
London
Aluminium (Al)
Calcium (Ca)
Lead (Pb)
Local Moran's I
GIS

ABSTRACT

Quantifying variation of metals in urban soils is crucial for efficient environmental and public health management. Prior to the interferences of human activities, regional topsoil geochemical distributions were generally primarily controlled by the underlying parent materials (PMs). With the rapid urbanisation, urban topsoil geochemical distributions have been modified with different levels. Three metals, aluminium (Al), calcium (Ca) and lead (Pb), were chosen to investigate the different levels of geogenic and anthropogenic controls in urban topsoil of Greater London Authority (GLA) area. These three metal elements clearly demonstrated the different spatial distribution affected by human activities using analysis of variance (ANOVA) and GIS-based spatial analysis. The high values of Al were in the clay which is located in the north and south GLA area; Pb accumulated in built-up areas and near traffic roads, which was likely associated with leaded paints and leaded petrol, respectively; the high Ca concentrations were in the Chalk bedrock of the southern GLA area and historical deconstruction and reconstruction sites in the city centre. Based on spatial interpolation and hot spot analysis (local Moran's I), this study reveals the different levels of geogenic and anthropogenic controls on different chemicals in urban soils: while the spatial distribution of Pb, which is more easily affected by human activities, can be significantly changed, the inert element Al may still be able to maintain its natural spatial distribution even in the heavily urbanized GLA area, and Ca demonstrates the mixed spatial distribution affected by both natural factors and human activities.

1. Introduction

Urban soils are more easily subjected to modification, especially by contamination because of various human activities (Wu et al., 2018). It is often believed that toxic metals, such as lead (Pb), cadmium (Cd) in urban areas are markedly affected by industrial activities and transport. However, parent materials (PMs) still have major natural control on soil geochemistry even in strongly urbanized areas (Argyraki and Kelepertzis, 2014). Because soils are derivatives of local bedrock through weathering, metals in soils are notably influenced by PMs' signatures (Eimil-Fraga et al., 2015; He et al., 2017). In this study, three metal elements were chosen to demonstrate the different levels affected by human activities in an urbanized environment. Identifying different impact levels of human activities on the different metals in topsoil helps to discover

* Corresponding author.
E-mail address: Chaosheng.Zhang@nuigalway.ie (C. Zhang).

https://doi.org/10.1016/j.chemosphere.2019.124541
0045-6535/

accumulation locations and potential sources.

As the capital of the United Kingdom, London is located in south-eastern England, with a population more than 8.67 million in 2015 (Lu et al., 2018). It has one of the largest urban zones in Europe, and urban soils are receiving more and more attention due to their potential association with public health (Li et al., 2018). In this study, three metal elements, aluminium (Al), calcium (Ca) and Pb, were selected to compare the extent of PMs' control and anthropogenic influences. Appleton et al. (2013) reported that the variance of Al, Ca and Pb in London region was controlled by soil PM dominantly, moderately and slightly, respectively, according to the results of analysis of variance (ANOVA). Aluminium is the most abundant metal in the crust where it comprises about 8% by weight. It is found as oxides or silicates in the mineral feldspars, beryl, and garnet and is largely controlled by PMs. Calcium is an essential nutrient element for plants and makes up the skeleton for animals. It forms minerals like calcite, dolomite, gypsum and fluorite and is influenced by both PMs and anthropogenic factors (e.g. water treatment facilities; destruction of buildings) (Appleton et al., 2013; Liu et al., 2015). Lead is recognised as one of the most common toxic heavy metals in city that poses magnificent harm to human health. The anthropogenic sources of Pb include coal combustion, steel works, sewage sludge, battery factories and remains of leaded painting (Laidlaw et al., 2017). It is elevated in the topsoil of urban landscape in Great Britain (Appleton et al., 2012). Children are more vulnerable to Pb poisoning than adults because of the most common pathway of hand-to-mouth behaviour. It is reported that Pb has a negative influence on intelligence development, which has been claimed to bring about an increase in rate of aggressive crimes (Taylor et al., 2016; Cave et al., 2018).

One-way ANOVA was utilised to compare the extent of geogenic and anthropogenic controls quantitatively, respectively. Parent materials represent the earth surface classification during natural (pre-human) period, while land use demonstrates the modern earth surface which is seriously modified by human impacts. Hence, two groups, PMs and land use, were chosen to explain geogenic and anthropogenic controls in this study. Furthermore, geographical information system (GIS) and spatial analyses, for instance, inverse distance weighted (IDW) interpolation and local Moran's I, are helpful to visualise the spatial distribution and investigate the hotspots. The local Moran's I, one of Local Indicators of Spatial Association (or Autocorrelation) (LISA) methods, has been widely used to identify spatial clusters and outliers of a variable by comparing to the values of their neighbouring samples (Zhang et al., 2008). The local Moran's I statistic is useful in the researches of air pollution (Zheng et al., 2016), soil contamination (Chen et al., 2016; Yuan et al., 2018), population growth (Salvacion and Magcale-Macandog, 2015), obesity rates (Huang et al., 2015), epidemiology (Dogru et al., 2017) and outlier detection (Fu et al., 2016).

The objectives of this study were (i) to visualise spatial distribution of Al, Ca and Pb in topsoil of Greater London Authority (GLA) area using spatial interpolation; (ii) to identify the different levels of influences of natural and anthropogenic factors on their spatial distributions using ANOVA and GIS-based spatial analyses; and (iii) to explore the natural and anthropogenic factors associated with the spatial patterns. The results of this study can be used to manage and control metal accumulation or contamination in soil.

2. Materials and methods

2.1. Soil parent materials

A simplified geological map (Miles and Appleton, 2005) was reproduced for the study (Fig. 1), showing a bedrock sequence from Cretaceous Period to Paleogene Period. The bedrock sequence of Cretaceous Period goes upwards from Lower Greensand Group, Selborne Group (Gault Formation and Upper Greensand Formation), Grey Chalk Group to White Chalk Group, while the bedrock sequence of Paleogene Period goes upwards from Lambeth group, Thames group to the top of Bracklesham group. Besides, some parts of the area are underlain by substantial Quaternary superficial deposits (British Geological Survey, 2011). The Quaternary deposits include clay-with-flints, plateau gravels, glacial till, river terrace deposits, brickearth, alluvium, and Head. The clay-with-flints were formed by weathering and solifluction of the original Palaeogene cover and dissolution of Chalk. Glacial till is deposited by the Anglian ice sheet which comprises mostly of pebbles or boulder-rich clay. River terrace deposits were formed by the diversion of the River Thames, and the gravel deposition crops out mainly on hilltops. Brickearth is loessic deposits with basal gravel originating from solifluction and fluvial. Alluvial deposits primarily form a flat surface in valleys of the rivers Thames and Lee, consisting principally of silty clay and clayey-silt with subsidiary sands.

2.2. Soil chemistry data

In the GLA area, a total of 6467 topsoil samples were collected using a hand-held auger by British Geology Survey (BGS) during 2005–2009. Although the sampling was done ten years ago, the potential hazard of accumulation metal, especially Pb, has been existing due to its low natural mobility (Li and Shuman, 1996). The sampling was taken to a depth of 5–20 cm on a 1 km × 1 km grid system. To minimize the influence of small-scale heterogeneity of the contaminants, each sample was a composite sample, consisting of 5 sub-samples, taken from 20 m × 20 m areas at the centre and four corners. After the drying process, soil samples were sieved to remove stones and plant debris through a 2 mm nylon mesh. Coarse powder was sealed into vessels containing mill balls after simply ground by hand in a pestle and mortar. After adding solid binder to improve adhesion of soil samples, milled powder was compressed into pellets (Johnson, 2011). Forty-eight trace and major chemical elements were determined by X-ray fluorescence (XRF) spectrometry, and loss on ignition (LOI at 450 °C) and pH were also measured. Details of sample preparation, analytical methods, and quality control procedures were described in Allen et al. (2011) and Johnson (2011).

2.3. Statistical and geostatistical analyses

The one-way ANOVA was applied to determine the proportion of the variance of element concentrations explained by PMs and land-use classes. ANOVA statistics are based on the assumptions: (i) the population has a normal distribution; (ii) observations are random and independent. But these requirements are seldom met entirely by geochemical data due to the demand of regular sample density. Logarithmic transformation (based on 10) was performed to improve the normality to meet the statistical assumption when the skewness coefficient exceeded 1.0 (Appleton et al., 2013).

GIS mapping and spatial analysis were used to show spatial distribution patterns and to identify the hotspots. The relatively simple interpolation method of IDW is sufficient to visualise the overall distribution pattern (Zhang, 2006), especially when the sample size is large. IDW assumes that each existing point datum has a local impact that decreases with distance. Power value and number of neighbours are the two main parameters for an IDW. The higher the power is, the stronger influence the nearest neighbour has on the prediction location. The number of neighbours to be included is to some extent arbitrary (Zhang et al., 2011).

The geochemical concentration hotspots can be recognised

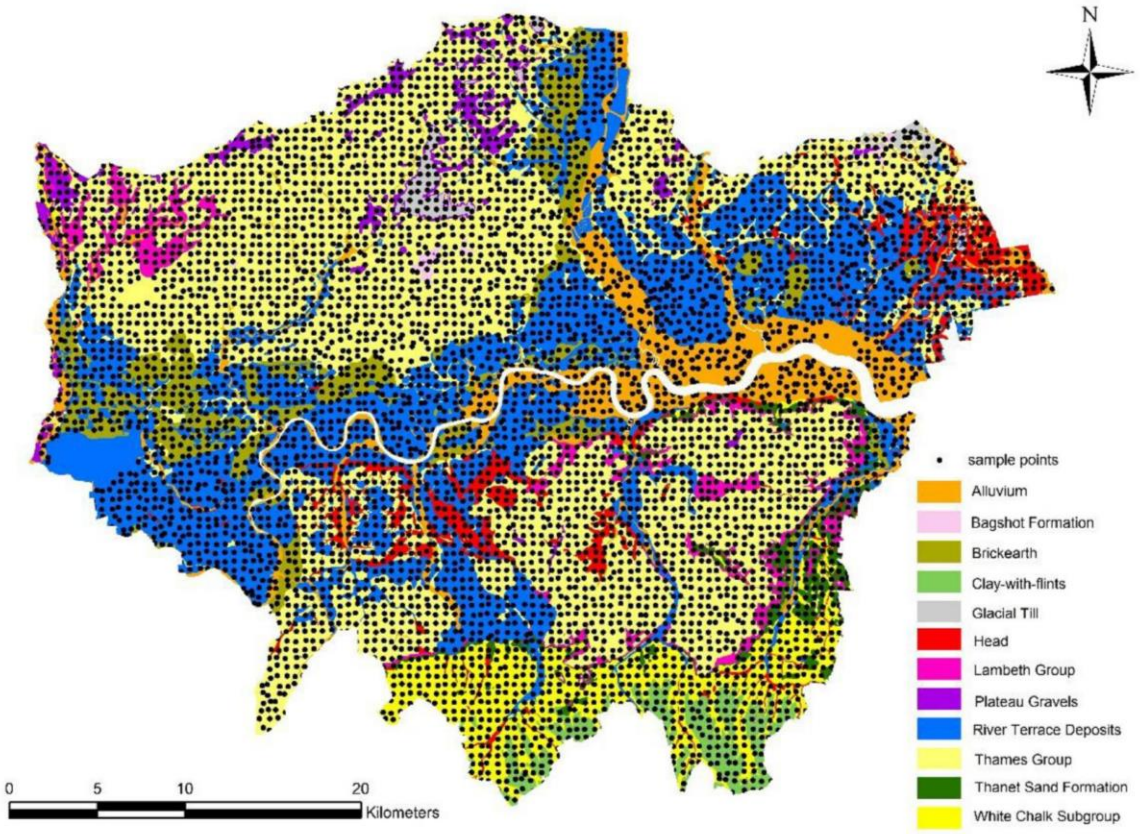


Fig. 1. Simplified geology map of the London region showing the boundary of the GLA area, overlaid with soil sampling locations.

through the local Moran's I index (Anselin, 1995; Getis and Ord, 1996):

$$I_i = \frac{z_i - \bar{z}}{\delta^2} \sum_{j=1, j\neq i}^{n} \left[w_{ij}(z_j - \bar{z}) \right] \quad (1)$$

where I_i denotes local Moran's I index; z_i is the value of the variable z (here Al, Ca or Pb) at location i; $\bar{z}$ is the mean value of z of the n samples; z_j is the value of the variable at all the other locations (where j ≠ i); σ^2 is the variance of variable z; and w_{ij} is a spatial weight which can be defined as fixed distance band or inverse distance between location pairs (Zhang et al., 2008).

The local Moran's I was carried out using 50 nearest neighbours as spatial weight with 9999 permutations to identify the location of clusters and outliers of the three metal concentrations in topsoil. It detects clusters as either "hot spots" (high values surrounded predominantly by high values, HH), or "cold spots" (low values surrounded predominantly by low values, LL), as well as "spatial outliers" as either high values surrounded primarily by low values (HL) or low values surrounded primarily by high values (LH) (Lalor and Zhang, 2001). The significance level of local Moran's I index can be assessed on a comparison of the calculated value with the values of the permutations (Anselin, 1995).

2.4. Data analyses and computer software

Raw data were stored in a file of MS Excel®. Basic statistical parameters were obtained and the test for normality was carried out using MINITAB® 15. The calculation of Local Moran's I index was performed in the software GeoDa (version 1.6.7 by Luc Anselin, 2015). All maps were produced using ArcGIS (version 10.3) software.

3. Results

3.1. Basic statistics of Al, Ca and Pb in urban soils of the GLA area

Table 1 demonstrates the descriptive statistics of Al, Ca and Pb. Al concentrations had an approximately normal distribution, while Ca and Pb concentrations had a strongly positively skewed distribution (Table 2 & Fig. 2). Hence, the log-transformed Ca and Pb data were used to perform the ANOVA and local Moran's I analysis. After data transformation, the data sets still could not pass the Kolmogorov-Smirnov test for distribution normality. However, the distribution was dramatically improved towards normality with smaller skewness values (Table 2 & Fig. 2). It has been found that the large sample size strengths the power of statistical test and leads to the rejection of normality test. Alternatively, graphical methods such as histogram, probability plots are helpful for a rough judgement for normality (Fig. 2) (Zhang et al., 2005).

3.2. ANOVA

The proportion of the variance (R^2 (adjusted)) in one-way ANOVA was applied to explain to what extent the PM controlled the geochemical concentrations. ANOVA results showed that there was a significant difference among PMs with R^2 (adj.) 26.2%, suggesting that PM exerts a strong control on Al (Table 3). Given that

Table 1
Summary statistics of Al, Ca and Pb for topsoil samples from the GLA area. n: total number of the samples; Min: minimum; Med: median; Max: maximum; DL: detection limit (concentrations in mg kg^{-1}).

Element	n	Min	5%	10%	25%	Med	75%	90%	95%	Max	Mean	DL
Al	6467	4235	22765	25941	31765	39705	51353	63000	69353	110118	42285	529
Ca	6467	1571	3071	4000	5786	9357	16571	29857	44042	376929	15923	357
Pb	6467	10.8	49.7	60.9	97.3	180.1	340	606.2	857.1	10000	295.6	1.2

Table 2
Skewness and kurtosis in the data sets with results of Kolmogorov-Smirnov test.

Element	Raw data			Log-transformed data		
	Skewness	Kurtosis	K–S *p*	Skewness	Kurtosis	K–S *p*
Al	0.63	0.13	<0.05			
Ca	6.25	53.85	<0.05	0.76	1.03	<0.05
Pb	8.81	144.76	<0.05	0.39	0.09	<0.05

Tukey HSD was helpful to determine specifically which type of PMs were different from each other, the results revealed that Thames Group and Clay-with-flints had higher concentrations of Al than the other types of PMs (Fig. 3 (a)). Thames Group is the most widespread of Palaeogene deposits across the GLA which is made of the thin Harwich Formation overlain by the thick London Clay Formation, cropping out in the north and central-south part of the GLA area. The London Clay Formation is typically composed of silty

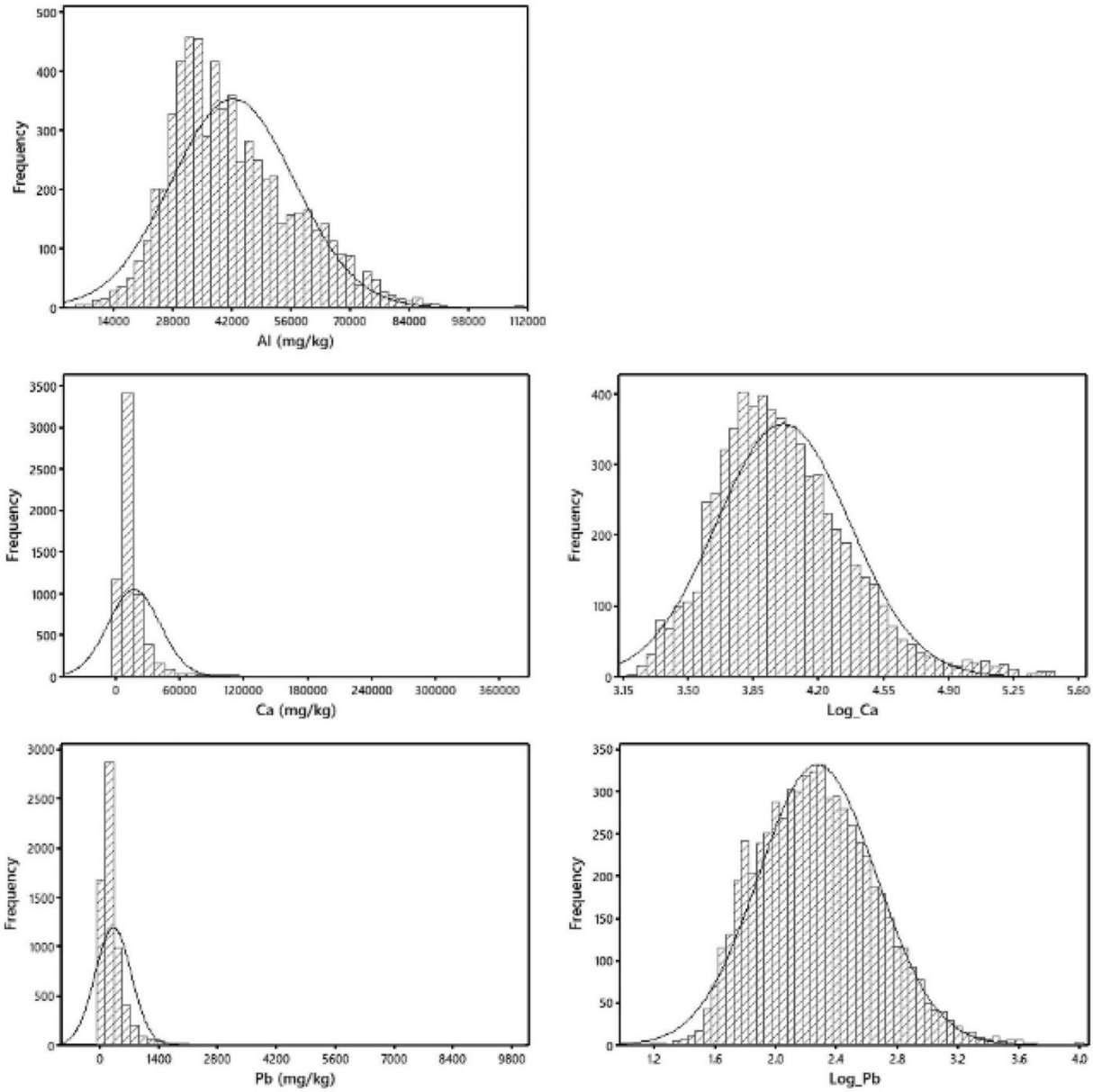


Fig. 2. Histograms for raw data of Al, Ca and Pb and log-transformed data of Ca and Pb (n = 6467). A normal distribution curve in histogram is superimposed for comparison.

Table 3
The results of ANOVA of Al, Ca and Pb in topsoil from the GLA area. n: number of the samples.

Element	n	DF	R^2 (adj)	*F*	*p*	Groups (sample number)
Al (PMs)	6467	11	26.2%	140.6	<0.001	Alluvium (558); Bagshot Formation (22); Brickearth (369); Clay-with-flints (159); Glacial Till (59);
Log_Ca (PMs)	6467	11	20.3%	149.6	<0.001	Head (291); Lambeth Group (177); Plateau Gravels (137); River Terrace Deposits (1605);
Log_Pb (PMs)	6467	11	8.2%	53.3	<0.001	Thames Group (2552); Thanet Sand Formation (124); White Chalk Subgroup (414)
Al (land use)	5268	8	3.6%	24.8	<0.001	Commercial (47); Residential (2936); Farm (424); Forest (408); Grass, Meadow & Heath (205);
Log_Ca (land use)	5268	8	8.0%	57.7	<0.001	Industrial (200); Other (278); Park & Nature Reserve (649); Recreation Ground (121)
Log_Pb (land use)	5268	8	17.3%	138.5	<0.001	

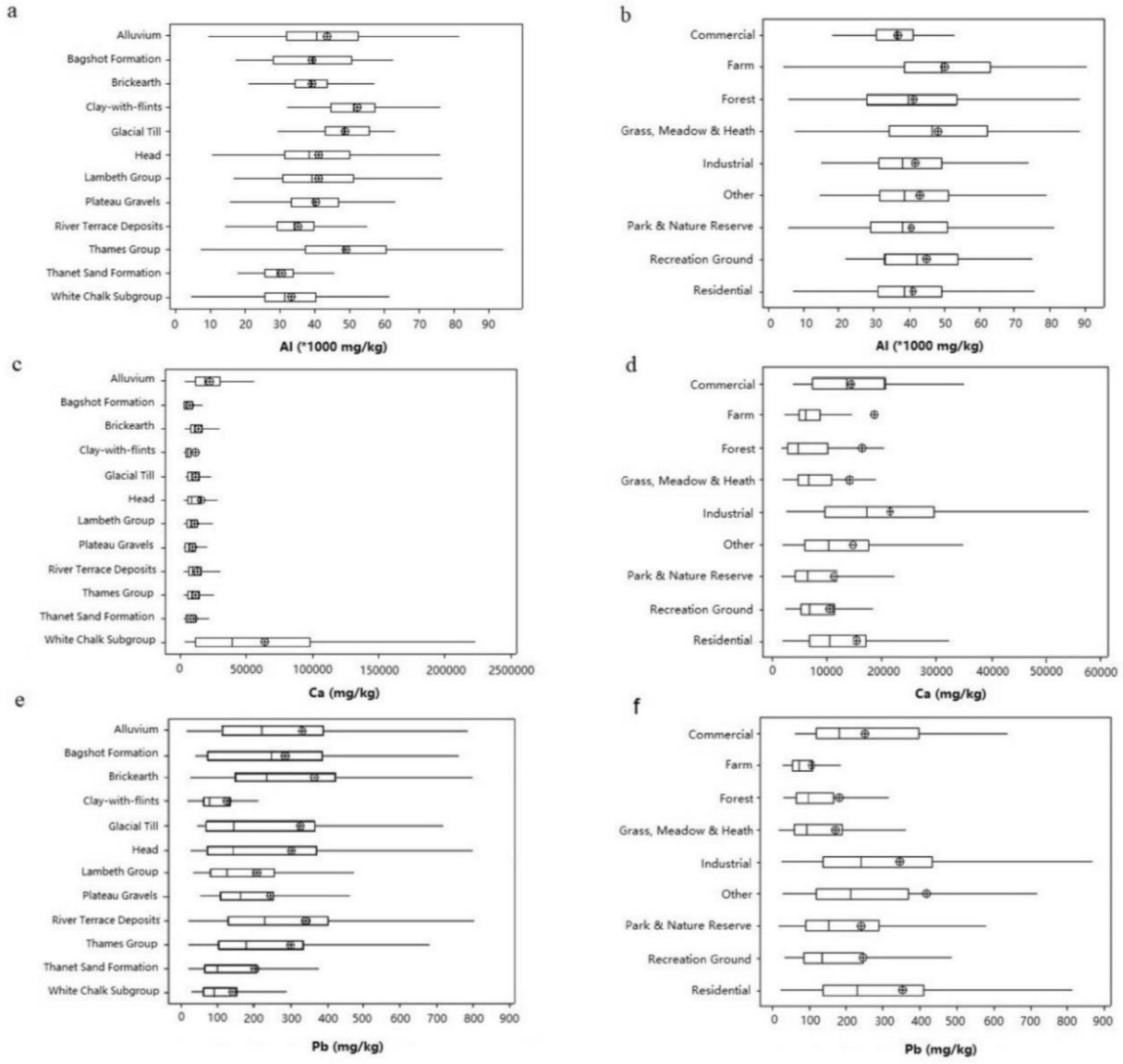


Fig. 3. Boxplot of Al, Ca and Pb concentrations in topsoil of different PMs groups (total number = 6467) and land use groups (total number = 5268) from the GLA area. Note: circle with cross = mean, box = interquartile range, whiskers extend to the lowest or highest value within the lower or upper limit, where lower limit = Q1 - 1.5(Q3-Q1) and upper limit = Q3 + 1.5(Q3-Q1).

clay and generally sandier towards the top of the Formation. Clay-with-flints is a residual deposit formed by weathering and solifluction of the original Palaeogene cover and earlier Quaternary deposits, cropping out in the south-east of GLA area. Clay consists of major components such as alumina (Al_2O_3), silica (SiO_2) in addition to oxides of iron (Fe_2O_3), titanium (TiO_2), and sodium (Na_2O) as minor components (Agbeleye et al., 2017). The result of ANOVA indicated that 20.3% of variance of Ca was explained by soil PM and

there was a significant difference among PMs (Table 3). The result of Tukey HSD suggested that White Chalk Subgroup and Alluvium had higher concentrations of Ca than the other types of PMs, White Chalk Subgroup in particular (Fig. 3 (c)). As it is well known, chalk is a soft, white, porous, sedimentary carbonate rock, a form of limestone composed of the mineral calcite. Calcite is an ionic salt made of calcium carbonate ($CaCO_3$). A low proportion (8.2%) of variance of Pb suggested that Pb was influenced by a minor degree of geogenic control. In spite of a significant difference among PMs according to the ANOVA analysis, the *F*-statistic value of Pb was much lower than *F*-statistic values of Al and Ca, indicating less significance of Pb than those of Al and Ca (Table 3). The result of Tukey HSD indicated that Brickearth had higher concentrations of Pb than other types of PMs, while White Chalk Subgroup and Clay-with-flints had lower concentrations of Pb than other types of PMs (Fig. 3 (e)).

The proportion of variance explained by land use for Al, Ca, Pb is 3.6%, 8.0%, 17.3%, respectively (Table 3). It suggested that land use exerted a weak control on Al. Farm and green land groups showed high Al concentrations, which scattered in the suburb area (Fig. 3 (b)). The topsoil overlaid on industrial, commercial and residential land-use areas had higher Ca concentrations than the rest types of land-use areas, according to the results of Tukey HSD (Fig. 3 (d)). Pb was elevated in industrial, commercial and residential land-use areas, verified by Tukey HSD analysis (Fig. 3 (f)).

3.3. Spatial distribution patterns (IDW and hot spot analyses)

The spatial distribution maps produced by IDW illustrated that high Al concentrations were located on the north, central-south and south-east parts of the GLA area, in agreement with the location of Thames Group and Clay-with-flints (Fig. 4 (a)). Fig. 4 (b) displayed that majority of hotspots overlapped two typical PMs, Thames Group and Clay-with-flints. Both IDW and hot spot maps confirmed that PMs have a great impact on Al spatial pattern in topsoil of the GLA area. Therefore, the soil mapping models based on PMs, even for urban geochemical mapping, were proven to be accessible (Appleton and Adlam, 2012). On the contrary, geomorphological change could indicate the type of underlying PMs as well (Tazikeh et al., 2017). As expected, Al tends to be high in clay-rich parent materials, such as Thames Group (clay), Clay-with-flints, and to a less extent Alluvium in the GLA area, while the relative low concentrations (cool spots) of Al were associated with gravels or sands PMs (e.g. River Terrace Deposits, Thanet Sand Formation) as well as the Chalk (Fig. 4 (b)). It is well known that the most common composition of sand in inland continental is silica (Si) or silicon dioxide (SiO_2), regardless of various constituents.

The high Ca values not only accumulated in the south part due to the chalk bedrock, but also gathered in the city centre of the GLA area (Fig. 4 (c) (d)). Chalk is a soft, white, porous, sedimentary carbonate rock. It is a form of limestone composed of the mineral calcite which is an ionic salt made of calcium carbonate ($CaCO_3$). The interesting elevation of Ca in topsoil across the centre was likely to be caused by anthropogenic influences. One reason of Ca enrichment in soils over alluvium on the banks of the River Thames and Lea, a known urbanized and industrial zone, maybe the massive use of debris from bomb-damaged buildings for land reclamation and flood prevention (Ward, 2015). Moreover, the drainage systems may bring the residual of industrial and residential waste, then enhance the Ca enrichment on Alluvium. Accumulation of Ca in city centre was also possibly linked by cement and lime-based mortar dust from the historical widespread destruction of buildings during the period 1940-41 as well as cement and concrete for reconstruction of new buildings since the beginning of 1940s in the urban domain (Appleton and Cave, 2018). Itkin et al. (2016) reported that calcrete cementation could be generated under the anthropogenic actions, aggravating the elevated $CaCO_3$ in topsoil. The relatively low concentrations (cool spots) of Ca in parks and national reserves, for instant Richmond park in southeast London, remarkably contrasted with hot spots in the built-up areas (Fig. 4 (d)).

The relatively high 'central zone' pattern of Pb concentration in the GLA area revealed by the IDW and hot spot map suggested that the Pb is notably influenced by anthropogenic activities (Fig. 4 (e) (f)). The hot spots of Pb gathered in the built-up area and the cool spots scattered in the non-built-up areas (Fig. 4 (f)). Table 4 displays Pb concentrations measured in soil in rural areas and urban soil of other UK cities. London had apparently high value of Pb compared by other regions which is associated with the level of urbanisation. The median value of Pb concentration of residential land-use (230 mg kg^{-1}) did not exceed the upper range of Category 4 Screening Levels (C4SLs) of Pb of residential land-use (330 mg kg^{-1}) (CL:AIRE, 2014), while the mean of Pb concentration of residential land-use (351 mg kg^{-1}) was higher than the upper range. C4SLs describe land that is clearly not contaminated land, with a low level of risk. The potential risk in the residential areas with high Pb is of great concern.

3.4. Exploration of anthropogenic Pb

To investigate the anthropogenic factors of Pb, the samples were divided into two groups, with sample number of 3246 and 3221, median values of 397.7 mg kg^{-1} and 192.7 mg kg^{-1}, in the built-up area and non-built-up area, respectively. Further comparisons between built-up areas and non-built-up areas were based on the data for the PMs with highest number of soil samples. They are Thames Group, River Terrace Deposits, Brickearth, White Chalk and Alluvium. Lead concentrations were remarkedly enriched in the built-up area soils overlaying the same PMs compared to the non-built-up area (Fig. 5). Built-up area mostly consists of industrial, commercial and residential areas in the GLA area, in agreement with Appleton and Cave (2018) (Fig. 3(f)). Old leaded paint originating from building surfaces is one of principal sources of Pb in soil in long-established built-up area (Hunt, 2016).

Except from urbanisation, the distribution of soil Pb may also be associated with road traffic (Johnson et al., 2017). Dong et al. (2017) reported that anthropogenic sources including recycled gasoline and road dust remained important for Pb in London. Burning gas resulted in release of Pb through vehicle emission, leading to Pb pollution to atmosphere, soil and crops (Xu and Liao, 2004). Pb in particles that can travel in air over a long distance from roads plays an important role in determining the concentrations of Pb in soil. Lin et al. (2000) described the Pb contamination of soil along a road as zonal distribution on both sides of the road up to 250 m. The primary road buffer (250 m) polygons were generated using the buffer tool in ArcGIS, according to the primary road shapefile of the GLA area downloaded from Geofabrik (2016). The divided sample sizes were 2056 and 4411, and the median values were 366.4 mg kg^{-1} and 262.6 mg kg^{-1} in the primary road buffer and out of primary road buffer area, respectively. The independent *t*-test was used to compare the difference in Pb concentrations between within primary road buffer and out of primary road buffer area after log transformation, and there was a significant difference between them ($p < 0.001$) for Pb concentration. Further contrasts between within-buffer areas and outside-buffer areas were based on the data for the PMs with highest number of soil samples, illustrating that the median values of Pb concentrations were notably higher in the within-buffer area soils overlaying the same PMs than the outside-buffer area (Fig. 6). Despite the cessation of utilization of leaded petrol since 2000, there was no discernible

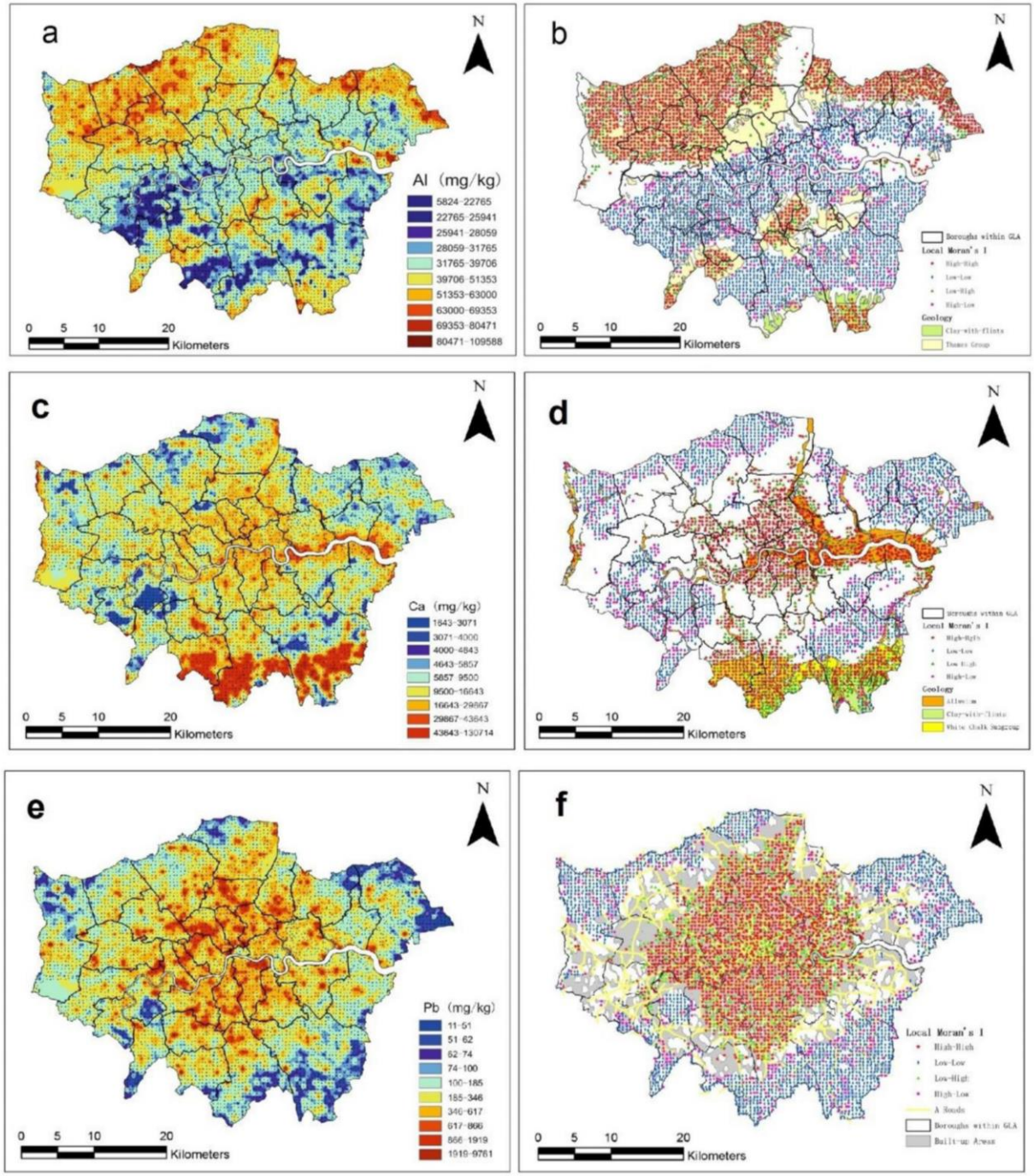


Fig. 4. IDW spatial interpolation map of (a) Al, (c) Ca and (e) Pb; Spatial distribution map of significant hotspots and cool spots for (b) Al and (d) Ca overlaying on main geology map; Spatial distribution map of significant hotspots and cool spots for (f) Pb overlaying on built-up map.

change in concentration of Pb in road surface soil between 2001 and 2010 (MacKinnon et al., 2011). The enrichment of Pb along the primary road was probably historical "traffic Pb". The similar finding of rise in Pb concertation with proximity to roads also reported in Appleton and Cave (2018). However, no substantial increase in iron (Fe), manganese (Mn) or Titanium (Ti) observed reduce the likelihood of significant vehicles contamination (Crosby et al., 2014). By contrast, low measured concentrations of Pb existed in several parks and national nature reserves located within the GLA area. These parks have not been disturbed by substantial residential, industrial activities and imported soil throughout recent 200–300 years. Generally, regardless of PMs, Pb spatial distribution was mainly influenced by anthropogenic control, such as built-up houses and buildings, leaded petrol from road traffic (Xue et al., 2017).

Table 4
Comparison of median Pb (mg kg^{-1}) in London urban topsoil with other urban, rural soils in England and Wales, European agricultural soils. Min: minimum; Max: maximum; n: number of the samples; NA: not available.

	Min	5%	10%	25%	50%	75%	90%	95%	Max	n	Reference
London	10.8	49.7	60.9	97.3	180.1	340	606.2	857.1	10000	6467	
Doncaster	20	29	32	43	69	118	256	405	3000	279	O' Donnell (2005a)
Lincoln	15	23	26	35	58	105	232	322	1400	216	O' Donnell (2005b)
Mansfield	1	19	25	43	76	141	249	322	1319	257	Freestone et al. (2004a)
Scunthorpe	9	19	21	31	45	76	149	219	3300	196	O' Donnell (2005c)
Sheffield	19	55	72	115	164	266	493	676	4300	575	Freestone et al. (2004b)
Rural areas in England and Wales	13	NA	33	39	49	73	134	NA	10000	5670	Rawlins et al. (2012)

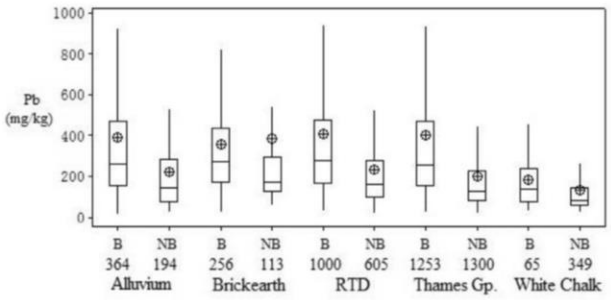


Fig. 5. Boxplot of Pb from five major PMs (RTD: River Terrace Deposits) subdivided into built-up (B) area and non-built-up (NB) area. Note: the number represents the sample number.

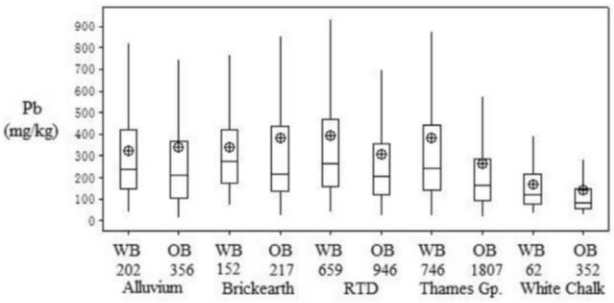


Fig. 6. Boxplot of Pb from five major PMs (RTD: River Terrace Deposits) subdivided into within-buffer (WB) area and outside-buffer (OB) area. Note: the number represents the sample number.

4. Conclusions

This study has demonstrated the different levels of three common metals in topsoil of the GLA area controlled by geogenic and anthropogenic: (i) The geogenic control on Al has not been destroyed even under the substantial human disturbance; (ii) Ca was controlled by a combination of geogenic and anthropogenic factors; (iii) Pb was primarily associated with anthropogenic activities. Naturally, Al tends to be enriched in clay material, and White Chalk Subgroup has high Ca concentration. In term of anthropogenic factors, Ca was closely related to building activities in the city centre. Pb had more complicated anthropogenic drivers, such as built-up, road traffic. Among the three metals, spatial distribution of Al is difficult to be modified by human activities, while that of Pb is widely affected in urban areas.

Acknowledgements

This study is sponsored by the Royal Society International Exchange Scheme of the UK (2015–2017). The soil data used in this study were provided by the BGS from the London Earth Project. Yuting Meng is funded by China Scholarship Council (No. 201504910793) for studying Ph.D.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2019.124541.

References

Agbeleye, A.A., Esezobor, D.E., Balogn, S.A., Agunsoye, J.O., Solis, J., Neville, A., 2017. Tribological properties of aluminium-clay composites for brake disc rotor applications. J. King Saud Univ. Sci. https://doi.org/10.1016/j.jksus.2017.09.002. (Accessed 8 September 2017).

Allen, M.A., Cave, M.R., Chenery, S.R.N., Gowing, C.J.B., Reeder, S., 2011. Sample preparation and inorganic analysis for urban geochemical survey soil and sediment samples. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 28–46.

Anselin, L., 1995. Local indicators of spatial association - LISA. Geogr. Anal. 27, 93–115.

Appleton, J.D., Adlam, K.A.M., 2012. Geogenic control on soil chemistry in urban areas: a novel method for urban geochemical mapping using parent material classified data. Appl. Geochem. 27, 161–170.

Appleton, J.D., Cave, M.R., 2018. Variation in soil chemistry related to different classes and eras of urbanisation in the London area. Appl. Geochem. 90, 13–24.

Appleton, J.D., Cave, M.R., Wragg, J., 2012. Modelling lead bioaccessibility in urban topsoils based on data from Glasgow, London, Northampton and Swansea, UK. Environ. Pollut. 171, 265–272.

Appleton, J.D., Johnson, C.C., Ander, E.L., Flight, D.M.A., 2013. Geogenic signatures detectable in topsoils of urban and rural domains in the London region, UK, using parent material classified data. Appl. Geochem. 39, 169–180.

Argyraki, A., Kelepertzis, E., 2014. Urban soil geochemistry in Athens, Greece: the importance of local geology in controlling the distribution of potentially harmful trace elements. Sci. Total Environ. 482–483, 366–377.

British Geological Survey, 2011. British geological survey official website. Available online at:, (Accessed 31 May 18). http://www.bgs.ac.uk/gbase/londonearth.html/.

Cave, M., Wragg, J., Lister, R., 2018. The effect of lead in soil on crime deprivation in Derby, Leicester and Nottingham. Appl. Geochem. 88, 198–212.

Chen, H., Lu, X., Gao, T., Chang, Y., 2016. Identifying hot-spots of metal contamination in campus dust of Xi'an, China. Int. J. Environ. Res. Pu. 13, 555.

CL:AIRE, 2014. Category 4 Screening levels (C4SLs). Available online at:, (Accessed 29 June 18). https://www.claire.co.uk/information-centre/water-and-land-library-wall/44-risk-assessment/207-category-4screening-levels-c4sls.

Crosby, C.J., Fullen, M.A., Booth, C.A., Searle, D.E., 2014. A dynamic approach to urban road deposited sediment pollution monitoring (Marylebone Road, London, UK). J. Appl. Geophys. 105, 10–20.

Dogru, A.O., David, R.M., Ulugtekin, N., Goksel, C., Zeker, D.Z., Sözen, S., 2017. GIS based spatial pattern analysis: children with Hepatitis A in Turkey. Environ. Res. 156, 349–357.

Dong, S., Gonzalez, R.O., Harrison, R.M., Green, D., North, R., Fowler, G., Weiss, D., 2017. Isotopic signatures suggest important contributions from recycled gasoline, road dust and non-exhaust traffic sources for copper, zinc and lead in PM10 in London, United Kingdom. Atmos. Environ. 165, 88–98.

Eimil-Fraga, C., Álvarez-Rodríguez, E., Rodríguez-Soalleiro, R., Fernández-Sanjurjo, M.J., 2015. Influence of parent material on the aluminium fractions in acidic soils under Pinus pinaster in Galicia (NW Spain). Geoderma 255–256, 50–57.

Freestone, S.E., O'Donnell, K.E., Brown, S.E., 2004a. Geochemical Baseline Data for the Urban Area of Mansfield. British Geological Survey Internal Report, p. 16. IR/02/082.

Freestone, S.E., O'Donnell, K.E., Brown, S.E., 2004b. Geochemical Baseline Data for

the Urban Area of Sheffield. British Geological Survey Internal Report, p. 14. IR/02/084.

Fu, W., Zhao, K., Zhang, C., Wu, J., Tunney, H., 2016. Outlier identification of soil phosphorus and its implication for spatial structure modelling. Precis. Agric. 17, 121–135.

Geofabrik, 2016. Geofabrik official website. Available online at:, (Accessed 15 May.18). http://download.geofabrik.de/europe/great-britain/england/greater-london.html.

Getis, A., Ord, J.K., 1996. Local spatial statistics: an overview. In: Longley, P., Batty, M. (Eds.), Spatial Analysis: Modelling in a GIS Environment. Geoinformation International, Cambridge, England, pp. 261–277.

He, S., Lu, Q., Li, W., Ren, Z., Zhou, Z., Feng, X., Zhang, Y., Li, Y., 2017. Factors controlling cadmium and lead activities in different parent material-derived soils from the Pearl River Basin. Chemosphere 182, 509–516.

Huang, R., Moudon, A.V., Cook, A.J., Drewnowski, A., 2015. The spatial clustering of obesity: does the built environment matter? J. Hum. Nutr. Diet. 28, 604–612.

Hunt, A., 2016. Relative bioaccessibility of Pb-based paint in soil. Environ. Geochem. Health 38, 1037–1050.

Itkin, D., Goldfus, H., Monger, H.C., 2016. Human induced calcretisation in anthropogenic soils and sediments: field observations and micromorphology in a Mediterranean climatic zone, Israel. Catena 146, 48–61.

Johnson, C.C., 2011. Understanding the quality of chemical data from the urban environment – Part 1: quality control procedures. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 61–76.

Johnson, L.E., Bishop, T.F.A., Birch, G.F., 2017. Modelling drivers and distribution of lead and zinc concentrations in soils of an urban catchment (Sydney estuary, Australia). Sci. Total Environ. 598, 168–178.

Laidlaw, M.A.S., Filippelli, G.M., Brown, S., Paz-Ferreiro, J., Reichman, S.M., Netherway, P., Truskewycz, A., Ball, A.S., Mielke, H.W., 2017. Case studies and evidence-based approaches to addressing urban soil lead contamination. Appl. Geochem. 83, 14–30.

Lalor, G., Zhang, C.S., 2001. Multivariate outlier detection and remediation in geochemical databases. Sci. Total Environ. 281, 99–109.

Li, Z., Shuman, L.M., 1996. Heavy metal movement in metal-contaminated soil profiles. Soil Sci. 161 (10), 656–666.

Li, G., Sun, G.X., Ren, Y., Luo, X.S., Zhu, Y.G., 2018. Urban soil and human health: a review. Eur. J. Soil Sci. 69, 196–215.

Lin, J., Qiu, Q., Chen, J., 2000. Assessment on pollution of heavy metals and metalloid in soil along road. J. Environ. Health 17 (5), 284–286 (in Chinese with English abstract).

Liu, H., Bu, H., Liu, G., Wang, Z., Liu, W., 2015. Effects of surrounding land use on metal accumulation in environments and submerged plants in subtropical ponds. Environ. Sci. Pollut. Res. 22, 18750–18758.

Lu, Q., Chang, N.-B., Joyce, J., Chen, A.S., Savic, D.A., Djordjevic, S., Fu, G., 2018. Exploring the potential climate change impact on urban growth in London by a cellular automata-based Markov chain model. Comput. Environ. Urban Syst. 68, 121–132.

MacKinnon, G., MacKenzie, A.B., Cook, G.T., Pulford, I.D., Duncan, H.J., Scott, E.M., 2011. Spatial and temporal variations in Pb concentrations and isotopic composition in road dust, farmland soil and vegetation in proximity to roads since cessation of use of leaded petrol in the UK. Sci. Total Environ. 409, 5010–5019.

Miles, J.C.H., Appleton, J.D., 2005. Mapping variation in radon potential both between and within geological units. J. Radiol. Prot. 25, 257–276.

O' Donnell, K.E., 2005a. Geochemical Baseline Data for the Urban Area of Doncaster. British Geological Survey Internal Report IR/02/079, p. 13.

O' Donnell, K.E., 2005b. Geochemical Baseline Data for the Urban Area of Lincoln. British Geological Survey Internal Report IR/02/081, p. 14.

O' Donnell, K.E., 2005c. Geochemical Baseline Data for the Urban Area of Scunthorpe. British Geological Survey Internal Report IR/02/083, p. 16.

Rawlins, B.G., McGrath, S.P., Scheib, A.J., Breward, N., Cave, M., Lister, T.R., Ingham, M., Gowing, C., Carter, S., 2012. The Advanced Soil Geochemical Atlas of England and Wales. British Geological Survey, Keyworth, Nottingham, pp. 134–135. Available online at: www.bgs.ac.uk/gbase/advsoilatlasEW.html/. (Accessed 13 August 2019).

Salvacion, A.R., Magcale-Macandog, D.B., 2015. Spatial analysis of human population distribution and growth in Marinduque Island, Philippines. Journal of Marine and Island Cultures 4 (1), 27–33.

Taylor, M.P., Forbes, M.K., Opeskin, B., Parr, N., Lanphear, B.P., 2016. The relationship between atmospheric lead emissions and aggressive crime: an ecological study. Environ. Health 15, 23.

Tazikeh, H., Khormali, F., Amini, A., Motlagh, M.B., Ayoubi, S., 2017. Soil-parent material relationship in a mountainous arid area of Kopet Dagh basin, North East Iran. Catena 152, 252–267.

Ward, L., 2015. The London County Council Bomb Damage Maps 1939-1045. Thames and Hudson, London.

Wu, W., Wu, P., Yang, F., Sun, D., Zhang, D., Zhou, Y., 2018. Assessment of heavy metal pollution and human health risks in urban soils around an electronics manufacturing facility. Sci. Total Environ. 630, 53–61.

Xu, P., Liao, C., 2004. Lead contamination of soil along road and its remediation. Chin. J. Geochem. 23 (4), 329–333.

Xue, S., Shi, L., Wu, C., Wu, H., Qin, Y., Pan, W., Hartley, W., Cui, M., 2017. Cadmium, lead, and arsenic contamination in paddy soils of a mining area and their exposure effects on human HEPG2 and keratinocyte cell-lines. Environ. Res. 156, 23–30.

Yuan, Y., Cave, M., Zhang, C., 2018. Using local Moran's I to identify contamination hotspots of rare earth elements in urban soils of London. Appl. Geochem. 88, 167–178.

Zhang, C., 2006. Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway, Ireland. Environ. Pollut. 142, 501–511.

Zhang, C., Manheim, F.T., Hinde, J., Grossman, J.N., 2005. Statistical characterization of a large geochemical database and effect of sample size. Appl. Geochem. 20, 1857–1874.

Zhang, C., Luo, L., Xu, W., Ledwith, V., 2008. Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. Sci. Total Environ. 398, 212–221.

Zhang, C., Tang, Y., Xu, X., Kiely, G., 2011. Towards spatial geochemical modelling: use of geographically weighted regression for mapping soil organic carbon contents in Ireland. Appl. Geochem. 26, 1239–1248.

Zheng, W., Li, X., Yin, L., Wang, Y., 2016. Spatiotemporal heterogeneity of urban air pollution in China based on spatial analysis. Rend. Fis. Acc. Lincei. 27, 351–356.

4.4 Comparison of methods for addressing the point-to-area data transformation to make data suitable for environmental, health and socio-economic studies

Meng, Y., Cave, M., Zhang, C., 2019. Comparison of methods for addressing the point-to-area data transformation to make data suitable for environmental, health and socio-economic studies. Science of the Total Environment 689, 797-807.

Summary:

Soil Pb is closely associated with human health as well as social behavior through soil ingestion and dust inhalation. We met the challenge of misaligned data when linking point Pb concentration to socioeconomic or healthy data of ward. The methodologies for integrating point and areal data have been reviewed. We conducted this article due to the need of considering the most suitable method. Among a number of methodologies, eight popular methods: (1) average, (2) median, (3) centroids IDW, (4) average block IDW, (5) median block IDW, (6) centroids OK, (7) average block OK and (8) median block OK, have been compared using Pb data set in the GLA area. The results indicated that the method of (5) median block IDW was recommended for further investigation in the ward-level of the GLA area. The reasons were (i) spatial interpolations were useful for estimating unobserved values when there was no samples collected in some wards; (ii) the median was more suitable than the average to represent a data set when it is heavily skewed; (iii) the block method decreased the estimation error and gave more representative values of ward than the centroid method; (iv) IDW preserved more spatial variation than OK, containing more local maxima (hot spots) and local minima (cool spots). Regardless of the difficulty in choosing the optimal method, this study has highlighted the point-to-area transformation issue and offered examples to compare the different methods.

My dedication in this paper accounted for 90% in reviewing literatures, exploring data and writing manuscript.

Science of the Total Environment 689 (2019) 797–807

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Comparison of methods for addressing the point-to-area data transformation to make data suitable for environmental, health and socio-economic studies

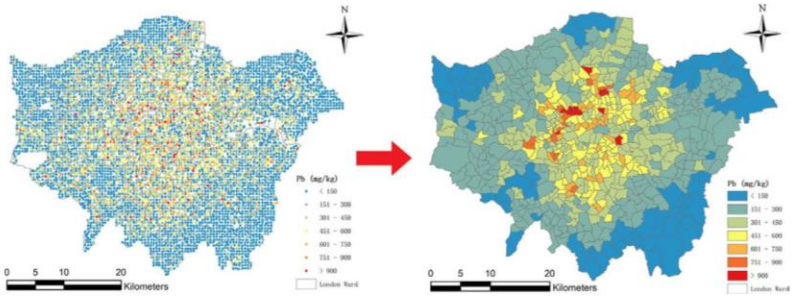
Yuting Meng [a], Mark Cave [b], Chaosheng Zhang [a,*]

[a] International Network for Environment and Health, School of Geography and Archaeology, Ryan Institute, National University of Ireland, Galway, Ireland
[b] British Geological Survey, Environmental Science Centre, Nottingham, United Kingdom

HIGHLIGHTS

- Methods for integrating point and areal data have been reviewed.
- The advantages and disadvantages of 8 methods were pointed out.
- Median block IDW is recommended to calculate Pb level in each ward of London.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 14 May 2019
Received in revised form 24 June 2019
Accepted 26 June 2019
Available online 27 June 2019

Editor: Jay Gan

Keywords:
GIS
Spatially-misaligned data
Point-to-area data transformation
Lead (Pb)
London

ABSTRACT

Soil lead (Pb) provides an important exposure pathway to the human body through soil ingestion and dust inhalation and is closely associated with human health as well as social behaviour. The challenge of transforming different spatial supports arises when linking point data (Pb concentration) to areal data (health status or social behaviour). A detailed review of methodologies for integrating point and areal data has been carried out. Among a number of methodologies, eight methods: (1) average, (2) median, (3) centroids inverse distance weighted (IDW), (4) average block IDW, (5) median block IDW, (6) centroids ordinary kriging (OK), (7) average block OK and (8) median block OK, have been compared using Pb data set in the Greater London Authority (GLA) area. The results indicated that the method of median block IDW was recommended for further investigation of the relationship between Pb concentration and socio-economic factors in the ward-level of the GLA area. The reasons were (i) spatial interpolations were useful for predicting unobserved values when simple average and median could not work in the locations where there were no samples collected in some areal units; (ii) the median value was more suitable than the average value for a skewed data set; (iii) the block method reduced estimation error and provided more representative values of areal units than the centroid method; (iv) IDW reserved more spatial variation than OK, containing more local maxima (hotspot) and local minima. Despite that it is still hard to decide the optimal method, this study has highlighted the point-to-area transformation issue and provided valuable examples to compare the different methods.

* Corresponding author.
E-mail address: Chaosheng.Zhang@nuigalway.ie (C. Zhang).

https://doi.org/10.1016/j.scitotenv.2019.06.452
0048-9697/

Table 1
Examples of methodologies for addressing the point-to-area data problem in environment, health and socio-economics research.

Variable A						Variable B						Reference
Name	Original		Transformation method	Transformed		Name	Original		Transformation method	Transformed		
	Type	Number		Type	Number		Type	Number		Type	Number	
Uranium (U) in topsoil	Point	n/a	Get the average value within each zip code region, Germany	Area	885	U in urine	Point	2253	Get the average value within each zip code region, Germany	Area	885	Schmidt et al., 2011
Pb in soil	Point	5467	Get the median value within each 1990 Census tracts, New Orleans, USA	Area	286	Pb in blood	Point	55,551	Get the median value within each 1990 Census tracts of New Orleans, United States	Area	286	Zahran et al., 2011
Blood Pb level	Point	59,645	Get the average value within each census tract	Area	106	Crime level	Area	106	No change	–	–	Boutwell et al., 2016
Pb in topsoil	Point	267 (Derby); 275 (Leicester); 284 (Nottingham)	Get the average value within each Lower Super Output Areas (LSOA) of Derby, Leicester and Nottingham, UK	Area	151 (Derby); 192 (Leicester); 182 (Nottingham)	The English Indices of Deprivation (EID)	Area	151 (Derby); 192 (Leicester); 182 (Nottingham)	No change	–	–	Cave et al., 2018
Arsenic (As) in soil	Point	270	Get the geometric mean value within each statistical local area (SLA), Victoria, Australia	Area	61	Cancer Incidence Rate	Area	61	No change	–	–	Pearce et al., 2012
As in drinking water	Point	124	(i) Use kriging interpolation; (ii) Get the predicted value at each maternal address, Arad, Bihor, and Timis, Romania	Point	294	As in urine	Point	294	No change	–	–	Neamtiu et al., 2015
As in groundwater	Point	1715	(i) Use an empirical model interpolation; (ii) Get town-level average value using zonal statistical tool of ArcGIS New Hampshire, USA	Area	n/a	Preterm birth rate; term low birth weight rate	Area	n/a	No change	–	–	Shi et al., 2015
As and chromium (Cr) in topsoil	Point	13,317	(i) Use kriging interpolation; (ii) Get the predicted value at each town's centroid, Spain	Area	7917	Cancer mortality	Area	7917	No change	–	–	Núñez et al., 2016
As and Cr in topsoil	Point	13,317	Use INLA+SPDE to predict the value at each town's centroid, Spain	Area	7917	Cancer mortality	Area	7917	No change	–	–	Núñez et al., 2016
Heavy metal in soil	Point	89	(i) use kriging interpolation; (ii) Convert the surface across Southern China into raster (150*150 m^2)	Area	n/a (raster resolution: 150*150 m^2)	Land use	Area	n/a (raster resolution: 10*10 m^2)	The land use in each upscale grid (150*150 m^2) was determined by the predominant land use in that grid.	Area	n/a (raster resolution: 150*150 m^2)	Zhao et al., 2012
Heavy metal in arable soil	Point	Over 388	Get the average value within each village, Shanxi, China	Area	97	Birth data and cases of birth defects	Area	97	No change	–	–	Zheng et al., 2012
Metal in topsoil	Point	1200	Get the median value within each area, South Carolina, USA	Area	10	Socioeconomic status	Area	10	No change	–	–	Davis et al., 2014
Heavy metal in soil	Point	2183	Collect samples in a 1∗1 km^2 grid, then convert it to raster data, Taiwan	Area	2183 (raster resolution: 1∗1 km^2)	Oral cancer mortality rates	Area	343	Use area-to-point Poisson Kriging to downscale the data from district-level to 1*1 km^2 grid scale	Area	2183 (raster resolution: 1∗1 km^2)	Lin et al., 2014
Metal in soil	Point	Approximately 992,000	(i) Use kriging interpolation; (ii) Get the predicted value at each maternal address, South Carolina, USA	Point	9920	Low birth weight	Point	9920	No change	–	–	McDermott et al., 2014
Heavy metal in topsoil	Point	1683	Get the average value within each town, Suzhou, China	Area	83	Cancer mortality rate	Area	83	No change	–	–	Chen et al., 2015
Heavy metal in farm soil	Point	n/a	Get the average value within each area, Taiwan	Area	12	Esophageal cancer rate	Area	12	No change	–	–	Lee et al., 2016
Heavy metal in farm soil	Point	n/a	Get the value of the proximity for each patient's address, Taiwan	Point	2725	Health state of male oral pre-malignant	Point	2725	No change	–	–	Tsai et al., 2017

Potentially toxic elements in topsoil	Point	13,860		Get the average value within each electoral ward, Northern Ireland, UK	Area	582	patients Cancer disease	Area	582	No change	–	–	McKinley et al., 2013
Geochemistry in groundwater and soil	Point	20,339 (groundwater); 10,738 (soil)		(i) Place a grid and use IDW interpolation; (ii) Get the average value of the predicted values at the nodes of the grid within each geological unit, Slovak Republic	Area	8	Health indicator	Area	8	No change	–	–	Fajčíková et al., 2014; Rapant et al., 2014
Chemistry in groundwater	Point	20,339		(i) Place a grid and use IDW interpolation; (ii) Get the average value of the predicted values at the nodes of the grid within each municipality, Slovak Republic	Area	2883	Mortality for cardiovascular diseases	Area	2883	No change	–	–	Rapant et al., 2015
Geochemistry in soil	Point	4856		Get the average value within each state, USA	Area	48	Disease mortality rates	Area	48	No change	–	–	Sun, 2018a, 2018b
Indoor radon	Point	5553		Get the average value within each district, Korea	Area	16	Lung cancer death	Area	16	No change	–	–	Lee et al., 2015
Air pollution exposure (O_3)	Point	56		(i) Use kriging interpolation; (ii) Get the predicted value at each Florida County's centroid, USA	Area	n/a	Myocardial infarction case	Area	n/a	No change	–	–	Young et al., 2009
Air pollution exposure (O_3)	Point	56		(i) Place a grid and use kriging interpolation; (ii) Get the average value of the predicted values at the nodes of the grid within each Florida County, USA	Area	n/a	Myocardial infarction case	Area	n/a	No change	–	–	Young et al., 2009
Air pollution exposure (SO_2, CO, NO_2, O_3, and PM_{10})	Point	13		Get the average value within Ulsan, Korea	Area	1	Lung function	Point	3827	Every participant was linked to the same average air pollution exposure			Son et al., 2010
Air pollution exposure (SO_2, CO, NO_2, O_3, and PM_{10})	Point	13		Get the value of the nearest monitor for each participant's address, Ulsan, Korea	Point	3827	Lung function	Point	3827	No change	–	–	Son et al., 2010
Air pollution exposure (SO_2, CO, NO_2, O_3, and PM_{10})	Point	13		(i) use IDW interpolation; (ii) Get the predicted value for each participant's address, Ulsan, Korea	Point	3827	Lung function	Point	3827	No change	–	–	Son et al., 2010
Air pollution exposure (SO_2, CO, NO_2, O_3, and PM_{10})	Point	13		(i) use kriging interpolation; (ii) Get the predicted value for each participant's address, Ulsan, Korea	Point	3827	Lung function	Point	3827	No change	–	–	Son et al., 2010
Air pollution exposure ($PM_{2.5}$, NO_2)	Point	78		Use a land use regression model to predict NO_2 for each residence; use a Eulerian dispersion model to predict $PM_{2.5}$ for each residence, Rome, Italy	Point	1,265,058	Health status of participant	Point	1,265,058	No change	–	–	Cesaroni et al., 2013
Air pollution exposure ($PM_{2.5}$, PM_{10}, O_3, CO, SO_2 and NO_2)	Point	Different station number for each pollutant		Get the average value within Montreal, Quebec, Canada	Area	1	Mortality among elderly persons	Area	1	No change	–	–	Goldberg et al., 2013
Air pollution exposure ($PM_{2.5}$ and PM_{10})	Point	n/a		Get the average value within each county, California, USA	Area	n/a	Breast cancer mortality	Area	n/a	No change	–	–	Hu et al., 2013

(continued on next page)

Table 1 (*continued*)

Variable A						Variable B					Reference
Name	Original		Transformation method	Transformed		Name	Original		Transformation method	Transformed	
	Type	Number		Type	Number		Type	Number		Type Number	
Air pollution exposure ($PM_{2.5}$ and O_3)	Point	n/a	(i) Use IDW interpolation; (ii) Get the predicted value at each participant's address, USA	Point	80,285	Lung cancer incidence	Point	80,285	No change	– –	Gharibvand et al., 2017
Air pollution exposure ($PM_{2.5}$, PM_{10} and NO_2)	Point	54	(i) use IDW interpolation; (ii) Convert the surface across Thailand into raster (iii) use GIS raster algebra analysis with the raster resampling technique for different resolutions	Area	40,676	Population distribution of exposure	Area	n/a	No change	– –	Pinichka et al., 2017
Air pollution exposure ($PM_{2.5}$ and O_3)	Point	n/a	(i) Use IDW interpolation; (ii) Get the predicted value at each zip code area's centroid, USA	Area	n/a	Mammographic breast density	Area	n/a	No change	– –	Yaghjyan et al., 2017
Air pollution exposure ($PM_{2.5}$)	Point	73	(i) use kriging interpolation; (ii) use zonal statistics function to calculate the average concentration within each township, Taiwan	Area	348	Bladder cancer mortality	Area	348	No change	– –	Yeh et al., 2017
Air pollution exposure ($PM_{2.5}$ and O_3)	Point	n/a	Get the average value within each spatial synoptic classification zone, Canada	Area	7	Mortality	Area	7	No change	– –	Cakmak et al., 2018
Air pollution exposure ($PM_{2.5}$, SO_2 and NO_2)	Point	8	Get the average value within Ningbo, China	Area	1	Ischemic heart disease mortality	Area	1	No change	– –	Huang et al., 2018
Air pollution exposure ($PM_{2.5}$)	Point	35	(i) Place a grid and use kriging interpolation; (ii) Get the average value of the predicted values at the nodes of the grid within each sub-district, Beijing, China	Area	325	Cancer mortality	Area	16	(i) assign the corresponding cancer mortality to the centroid of each district; (ii) Place a grid and use kriging interpolation; (iii) Get the average value of the predicted values at the nodes of the grid within each sub-district	Area 325	Ouyang et al., 2018
Air pollution exposure ($PM_{2.5}$)	Point	1082	(i) Place a grid and use IDW interpolation; (ii) Get the average value of the predicted values at the nodes of the grid within each county, USA	Area	n/a	Hepatocellular carcinoma incidence	Area	n/a	No change	– –	VoPham et al., 2018
Air pollution exposure ($PM_{2.5}$)	Point	n/a	(i) Place a grid and use specific model; (ii) Get the average value of the predicted values at the nodes of the grid within each region, China	Area	103	Pancreatic cancer mortality	Area	103	No change	– –	Wang et al., 2018
Self-reported health status	Point	4131	No change	–	–	Socio-economic characteristics	Area	n/a	(i) Place a buffer zone around each participant; (ii) Get the area-weighted average characteristics within each buffer zone, Amsterdam, the Netherlands	Area 4131	Veldhuizen et al., 2013

1. Introduction

As the environment has significant impact on the health status, especially in heavily contaminated or geologically unfavourable areas, numerous scientific articles have reported the associations between the environmental components (e.g. geochemistry, air pollution) and human health (Rapant et al., 2014). Nevertheless, a number of studies have found that lead (Pb) exposure was potentially related to health (Veldhuizen et al., 2013; Davis et al., 2014; Farkhondeh et al., 2015; Carlson and Neitzel, 2018); socioeconomic status (Gump et al., 2009), socio-behaviour (Naicker et al., 2012), and even crime (Mielke and Zahran, 2012; Feigenbaum and Muller, 2016; Cave et al., 2018). However, there is a challenge in the use of different spatial scale data when more and more interdisciplinary researches focus on the relationships among environment, health and socio-economics. For instance, environmental data, including soils (e.g. concentration of heavy metal) and air pollution (e.g. air monitoring station), are often collected as point data; health data (e.g. mortality rate, cancer morbidity rate) are usually aggregated in administration areal unit level; socio-economic data (e.g. education, ethnicity, income) are commonly recorded by census or survey in designed regions in order to ensure data confidentiality. The different types of spatial data (point or area) cause a problem in the process of synthesis of spatial data (Gotway and Young, 2002; Young and Gotway, 2007). However, the performances of the transformation methods have not been fully assessed and it deserves attention to make sure that the transformation is reasonable and appropriate for specific studies.

Table 1 delineates the methodologies which were used to solve the problem of point-to-area transformation in some recent articles. Most environmental health studies have chosen simple approaches to estimate element concentrations or exposure from air pollutants on the spatially aggregated level, such as using a value from one nearest sample or monitoring station (Son et al., 2010; Tsai et al., 2017), average values of samples or monitoring stations within each areal unit (Son et al., 2010; Schmidt et al., 2011; Goldberg et al., 2013; Hu et al., 2013; McKinley et al., 2013; Chen et al., 2015; Lee et al., 2015; Boutwell et al., 2016; Lee et al., 2016; Cakmak et al., 2018; Cave et al., 2018; Huang et al., 2018; Sun, 2018a; Sun, 2018b). One shortcoming of these methods is neglecting spatial heterogeneity which may cause inaccuracy in estimating the influences of the environment. In order to minimize the aggregation effect and increase the precision of estimation, geographic information system (GIS) techniques are widely used in integrating incompatible spatial data. For example, inverse distance weighted (IDW) interpolation (Son et al., 2010; Gharibvand et al., 2017; Pinichka et al., 2017), kriging interpolation (Young et al., 2009; Zhao et al., 2012; Son et al., 2010; McDermott et al., 2014; Neamtiu et al., 2015; Núñez et al., 2016) and block kriging method (Young et al., 2009; Fajčíková et al., 2014; Rapant et al., 2014; Rapant et al., 2015) have been applied in many studies. Moreover, air quality modelling and empirical modelling were used to predict environmental indices (Cesaroni et al., 2013; Shi et al., 2015). Besides, Bayesian modelling (the integrated nested laplace approximation (INLA) combined with a stochastic partial differential equations (SPDE)) has been selected in modelling and mapping environmental variables (Núñez et al., 2016; Poggio et al., 2016; Huang et al., 2017). While these geostatistical procedures which have been applied to the data follow the upscaling direction, area-to-point kriging was also developed for downscaling areal data to point support (Lin et al., 2014; Wang et al., 2016).

Lead is one of the toxic metal elements that has recently received considerable attention. Even low-level exposure to Pb can cause adverse health effects (Markus and McBratney, 2001). Soil is an important pathway of human Pb exposure which more often occurred in children via hand-to-mouth transportation (Sahmel et al., 2015). The highest Pb contamination in soil was more likely to occur in city centres than suburbs (Mielke and Reagan, 1998). Appleton and Cave (2018) have shown a relatively high 'central zone' pattern of Pb concentration in the Greater London Authority (GLA) area. This pattern has higher risk due to high population density in the city centre. The elevated level appeared to be caused by destruction of buildings and road traffic. Lark and Scheib (2013) have reported that industrial sites had the largest mean Pb content, followed by domestic gardens. Eating vegetables that grew in the domestic gardens is another pathway to transfer soil Pb to human beings (Sung and Park, 2018). Therefore, it is of great importance to investigate the association between Pb in soil and socioeconomic or health status in the GLA area. However, we need to solve the problem of transformation from point data to areal data before this can be carried out, since different transformation methods are likely to affect the result.

In this study, we attempted to transform point-level Pb concentrations to ward-level estimates in order to join with the ward-level socio-economic data. There are different levels of statistical units in the GLA area, such as borough, ward, middle super output areas, lower super output areas. The ward level was selected because it has a medium size and it is a common administration unit. The aim of this research was to compare the basic methods of estimating concentration of soil Pb. We applied eight different common methods: (1) average (the average value within each ward); (2) median (the median value within each ward); (3) centroids IDW (the predicted value on the centroid within each ward using IDW interpolation); (4) average block IDW (to place a grid and get the average value of the predicted values at the nodes of the grid within each ward using IDW interpolation); (5) median block IDW (to place a grid and get the median value of the predicted values at the nodes of the grid within each ward using IDW interpolation); (6) centroids ordinary kriging (OK) (the predicted value on the centroid within each ward using OK interpolation); (7) average block OK (to place a grid and get the average value of the predicted values at the nodes of the grid within each ward using OK interpolation); (8) median block OK (to place a grid and get the median value of the predicted values at the nodes of the grid within each ward using OK interpolation). The methods of (1) average and (2) median are the most straightforward way for estimation over large areas that contain randomly located observations, while the rest of methods are used for prediction from known data which are based on two commonly used interpolation methods of IDW and OK. Geostatistical approaches have been widely used to assess spatial structure and to investigate spatial distribution of soil heavy metals (Huo et al., 2012; Shaheen and Iqbal, 2018). IDW is easy to handle due to fewer input parameters, while kriging is often preferred which is an unbiased estimator but more complicated to use (Qin et al., 2015). Except for the difference of average and median calculation, the significant differences between the centroid method and average or median methods are investigated. According to the same spatial interpolation, the centroid method adopts one predicted single value at the central point of a

Table 2
The prediction accuracy of different interpolation methods.

Interpolation method	IDW	IDW	IDW	IDW (selected)	IDW	IDW	OK	OK	OK (selected)
Data transformation	None	None	None	None	None	None	Logarithm	Logarithm	Logarithm
Number of neighbours or model	5	10	15	20	30	50	Spherical	Gaussian	Exponential
MRE	0.93	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95
RMSE	438.13	421.70	415.68	412.94	412.33	412.02	407.78	407.76	403.10
R^2	0.07	0.09	0.10	0.10	0.10	0.10	0.11	0.11	0.12

Table 3
Summary statistics of Pb (mg kg^{-1}) for topsoil samples from the GLA area. TN: total number of the samples; SD: standard deviation; CV: coefficient of variation; Min: minimum; Med: median; Max: maximum; DL: detection limit.

TN	Mean	SD	CV	Min	5%	10%	25%	Med	75%	90%	95%	Max	Skew	DL
6467	295.6	430.4	1.46	10.8	49.7	60.9	97.3	180.1	340	606.2	857.1	10,000	8.81	1.2

polygon to represent one area, while average or median methods use average or median value to describe each area.

2. Materials and methods

2.1. Study area

London, the study area, is the capital of the United Kingdom which has become a major human settlement since two thousand years ago. The population was 8.2 million based on the 2011 Census and would continue to grow in the next few decades (Great London Authority, 2016). With the intensive urbanization, the land use and soil quality could change dramatically in the urban areas. There are 654 wards within the GLA area. The GIS shapefile used in this study has 630 wards because 25 wards within "City of London" have been merged into a single area. The reason is that the wards within "City of London" are relatively smaller compared to other wards outside City of London (National Statistics data, 2015). The ward is the primary unit of English electoral geography for civil parishes and borough and district councils.

2.2. Soil chemistry data

A total of 6467 topsoil (ca. 5–20 cm depth) samples were collected by British Geology Survey (BGS) using a hand-held auger in the GLA area. At each site, composite samples based on 5 sub-samples were taken from a 20 m × 20 m area at the centre and four corners in order to minimize the influence of small-scale heterogeneity of the contaminants. Forty-eight trace and major chemical elements (< 2 mm size fraction) were determined by X-ray fluorescence (XRF) spectrometry, together with loss on ignition and pH. More details of sample preparation, analytical methods, and quality control procedures have been previously described in Allen et al. (2011) and Johnson (2011).

2.3. Spatial data analysis

2.3.1. Inverse distance weighted

Inverse distance weighted interpolation is one of a number of deterministic interpolation methods (Shepard, 1968). It assumes that measurements that are close to one another are more similar than those that are farther apart. As a result, the measured values that are closer to the predicted point are assigned greater weights than those farther away (Franke and Neilson, 1980). In other words, each value of point of interest is a distance weighted average of nearby measured observations (Babak and Deutsch, 2009). There are two major parameters for an IDW, the power value and the number of neighbours (Zhang et al., 2011). However, there is no criterion to decide the optimal parameters. In this study, the power value that decides how rapidly the influence of the point diminishes with distance was determined as 2 which is most commonly used. The number of neighbours was chosen as 20. Because the cross-validation showed a relatively small value of root mean square error (RMSE) and the spatial distribution surface contained a more detailed pattern, in comparison to 5, 10, 15, 30, 50 neighbours (Table 2).

2.3.2. Kriging

Kriging is also a weighted combination of observed values with the total weights summing to 1. However, it is based on spatial autocorrelation among the observed points rather than assuming a function of inverse distance. Because of this, kriging can provide unbiased prediction with minimum variance. Before implementing a kriging, a model should be selected to fit the semivariogram (Oliver and Webster, 2015). Ordinary kriging, one of the most commonly used types in the kriging family, was applied in this study. OK has ability to create map smoothly with small standard errors (Wang et al., 2014). In an OK, the weight is determined by the fitted model to the observed points, the distance to the prediction location, and the spatial relationships among the observed values surrounding the prediction location (Krige, 1966). The exponential model with the lowest RMSE and highest R^2 was selected (Table 2) (Robertson, 1995).

2.3.3. Block kriging

Block kriging is a routine method to solve the point-to-area transformation in term of geostatistics. Unlike the centroid approach that predicts every ward at point-level, block kriging estimates spatial averages for each ward. To carry out block kriging, a fine grid (100 m × 100 m) was placed across the GLA area, then OK was applied to estimate Pb concentration at each node on the grid. Finally, the estimated Pb concentrations within a ward were averaged to represent the estimated Pb concentration for that ward (Young et al., 2009). It is necessary to point out that the size, shape and orientation of the chosen block (grid) would affect the result of estimating the average values (Gotway and Young, 2002). The mathematic details of this methodology were presented in previous studies (Journel and Huijbregts, 1978; Cressie, 1993; Gotway and Young, 2002). Block IDW is similar with block kriging which uses IDW interpolation instead of OK. In this study, the block OK and block IDW are based on the selected OK and IDW interpolation, respectively.

2.4. Comparison of the accuracy of the interpolation methods

We used three parameters to compare the accuracy of the interpolation methods. The mean relative error (MRE) and the RMSE were calculated from the observed and estimated values at each sample site.

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{O_i - E_i}{O_i}\right|$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[O_i - E_i]^2}$$

where O_i is the observed value at location i, E_i is the estimated value at location i, and n is the sample size. The smaller MRE and RMSE is, the higher accuracy of the interpolation is.

The correlation coefficient (R^2) is another criterion that is an assessment of the correlation between the measured and predicted values (Yao et al., 2014). It can be calculated using the following equation:

$$R^2 = \frac{\left[\sum_{i=1}^{n}(O_i - O_{ave})(E_i - E_{ave})\right]^2}{\sum_{i=1}^{n}(O_i - O_{ave})^2 \sum_{i=1}^{n}(E_i - E_{ave})^2}$$

where O_{ave} is the average of the measured values, E_{ave} is the average of the estimated values.

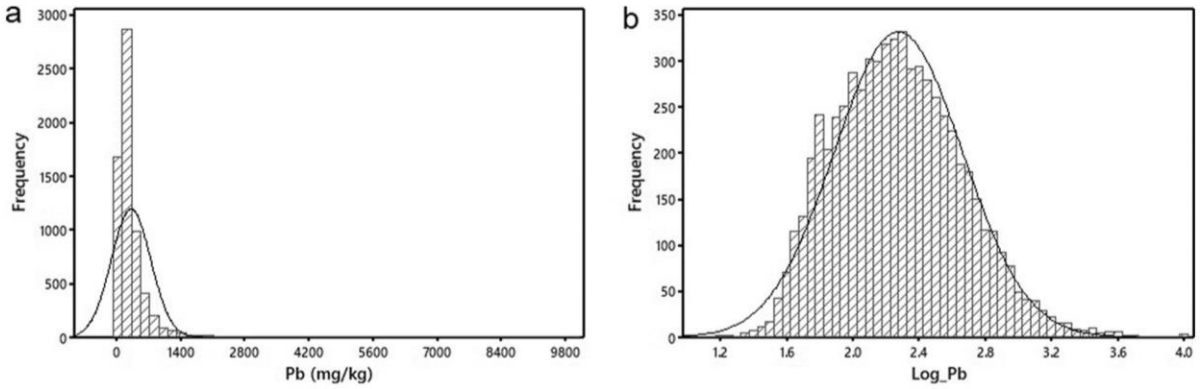


Fig. 1. Histogram of Pb for topsoil samples from the GLA area. (a) Raw data; (b) Log-transformed data.

2.5. Data analyses and computer software

Raw data were stored in a file of MS Excel®. Summary statistical parameters were obtained using SPSS® ver. 23. All maps were produced using ArcGIS (ver. 10.3) software.

3. Results

3.1. Descriptive statistics of Pb data set

Table 3 demonstrates that Pb concentrations had a strong positively skewed distribution (Fig. 1(a)). The untransformed data was used for IDW interpolation, while the log-transformed data was used for kriging interpolation (Fig. 1(b)). Fig. 2 presents the Pb concentrations of sampling sites in the GLA area. The mean and median values of Pb concentration, 295.6 mg kg^{-1} and 180.1 mg kg^{-1} respectively, were remarkedly beyond the safety limits (80 mg kg^{-1}) in soil (Mielke et al., 2000).

3.2. Spatial distribution of ward-level Pb

All ward-level estimated maps displayed a similar pattern of Pb distribution that high concentrations were clustered in the city center (Fig. 3). The spatial pattern indicated that the city center is strongly disturbed by human activities. However, each map produced by different method showed comparable, but different details of Pb distribution. After changing the support of the original variable, the new variable has different statistical and spatial properties.

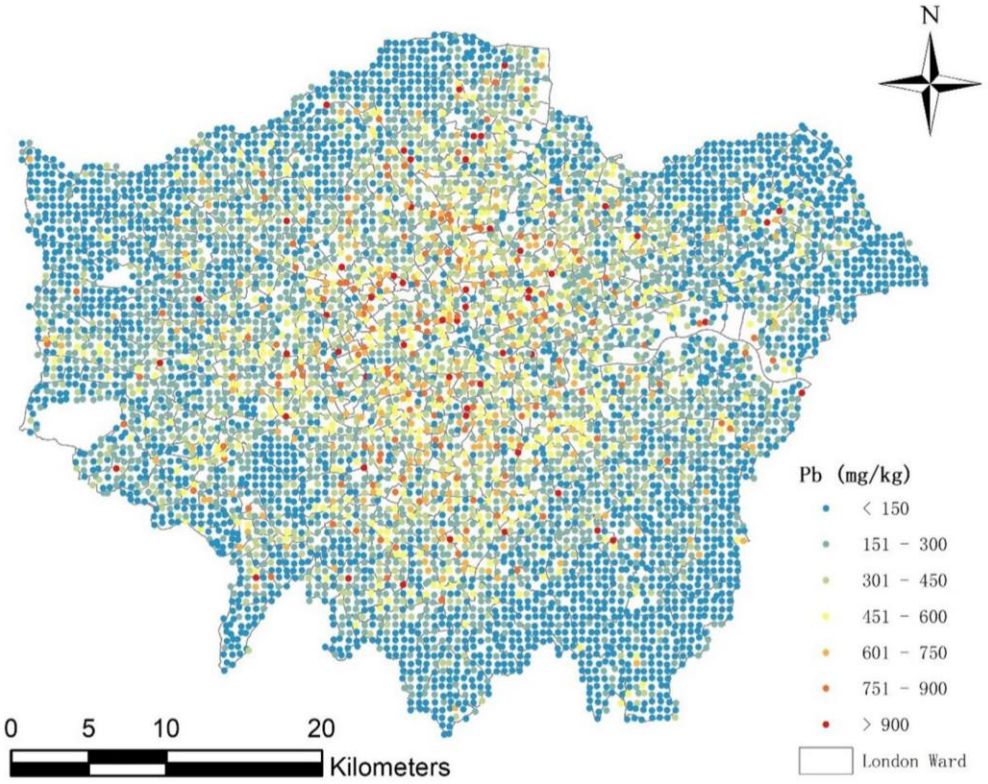


Fig. 2. The point Pb concentration distribution map.

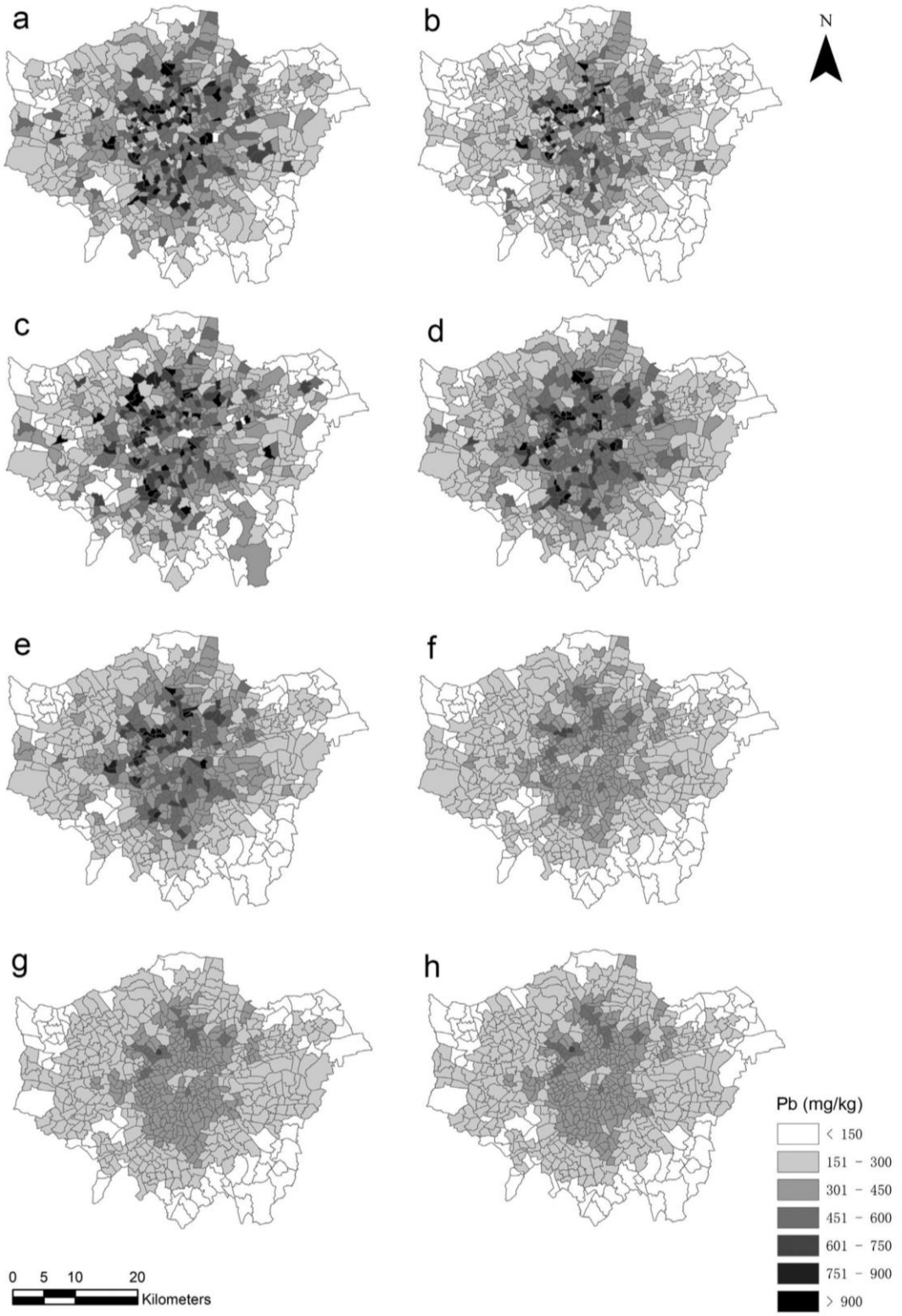


Fig. 3. Predicted ward-level topsoil Pb concentration (mg kg^{-1}) for the GLA area by using (a) average, (b) median, (c) centroids IDW, (d) average block IDW, (e) median block IDW, (f) centroids OK, (g) average block OK and (h) median block OK.

Table 4
Summary statistics of estimated ward-level Pb concentration for the GLA area from eight different methods. N: number of the samples; SD: standard deviation; CV: coefficient of variation; Min: minimum; Med: median; Max: maximum.

Name	N	Type	Min	25%	Median	75%	Max	Mean	SD	CV
Topsoil samples	6467	Point	10.8	97.3	180.1	340	10,000	295.6	430.4	1.46
Average	629	Area	71.4	218.1	322.6	498.4	2152.2	393.2	265	0.67
Median	629	Area	57.6	169.4	257.9	386.7	1588.5	311.0	217	0.70
Centroid IDW	630	Point value represents area	42.0	195.6	335.7	486.7	3936.3	385.5	301	0.78
Average block IDW	630	Area	76.6	238.3	351.7	493.3	1438.9	390.9	208.5	0.53
Median block IDW	630	Area	70.2	215.2	322.9	459.4	1226.2	355.1	184.8	0.52
Centroid OK	630	Point value represents area	59.8	187.9	266.1	351.7	665.5	275.7	110.8	0.40
Average block OK	630	Area	74.2	191.4	266.5	356.3	643.1	275.3	105.4	0.38
Median block OK	630	Area	72.7	188.2	265.4	351.2	650.5	273.0	106.2	0.39

There were 630 wards within the GLA area with an average area of 2.5 km^2. The sampling density is relatively high in this study, however, there was still one ward (Bayswater) without any soil sampling point, so the average and median values of this ward were missing (Fig. 3 (a), (b)). Although methods of (1) average and (2) median are the simplest way to estimate the mean values, they are not suitable with sparse sampling or uneven random sampling.

Mean values of methods of (1) average and (4) average block IDW were higher than the rest of the approaches (Table 4). This is because the average is influenced by outliers which are much higher or lower than the rest of the values. In this case, the median was a resistant midpoint measurement to filter out outliers where the data set was skewed.

The interpolations methods were employed to solve the problem of lack of sampling in some areal units. The centroid IDW approach provided the largest standard deviation (SD) and coefficient of variation (CV) (Table 4). In statistics, SD and CV are used to quantify the amount of variation of a data set. The reason behind this is that the variability in observation on point-level supports is considerably more than the variability in statistical averages on areal units. Spatial prediction at areal centroids could delineate a misleading picture of geographic variation in an areal variable because variables associated with areal units are inherently spatial averages (Young et al., 2009).

Comparing maps (Fig. 3 (d), (e), (g), (h)) to maps (Fig. 2 (c), (f)), the predicted values from methods ((4) average block IDW, (5) median block IDW, (7) average block OK and (8) median block OK) appeared smoother than those from methods ((3) centroids IDW and (6) centroids OK). The centroid methods only rely on prediction at one point which is easily biased. The block methods are based on the fundamental central limit theorem which extract an average to predict a variable in an area. There was a greater tendency of wards close together to have predicted Pb values more alike than wards further apart for the block approaches than for the centroid method, which is analogous to the First Law of Geography (Tobler, 1970). The "block" method has been generalized, thus it can be used for aggregation, disaggregation, side-scaling and intensity estimation (Young et al., 2009). In general, the block interpolation method produces more reliable predicted values when the sampling density is low. However, the alternative spatial prediction technique also has disadvantage which is the uncertainty and estimated errors with predicted values.

To summarize, methods of (5) median block IDW and (8) median block OK would be selected for further consideration since (i) surface interpolations were a powerful tool for predicting unobserved values; (ii) the median value was more suitable than the average value for skewed data set; (iii) the block method reduced estimation variance and provided more representative values of areal units than the centroid method.

3.3. Accuracy assessment of interpolation methods

Table 2 summarized the values of MRE, RMSE and R^2 of cross validation. The results indicated that the accuracy of OK was slightly better than that of IDW. However, compared with the surface maps based on two interpolation methods, we could discern that map produced using OK was much smoother than the one produced using IDW (Fig. 4). This "smoothing effect" is a common drawback of the kriging method in which small values are often overestimated and large values are underestimated. A multifractal concept can be adopted in both IDW and Kriging interpolation which incorporates local singularity into the basic model in order to address the smoothing effect issue (Cheng, 2001; Yuan et al., 2012). Moreover, bigger size of sample would increase

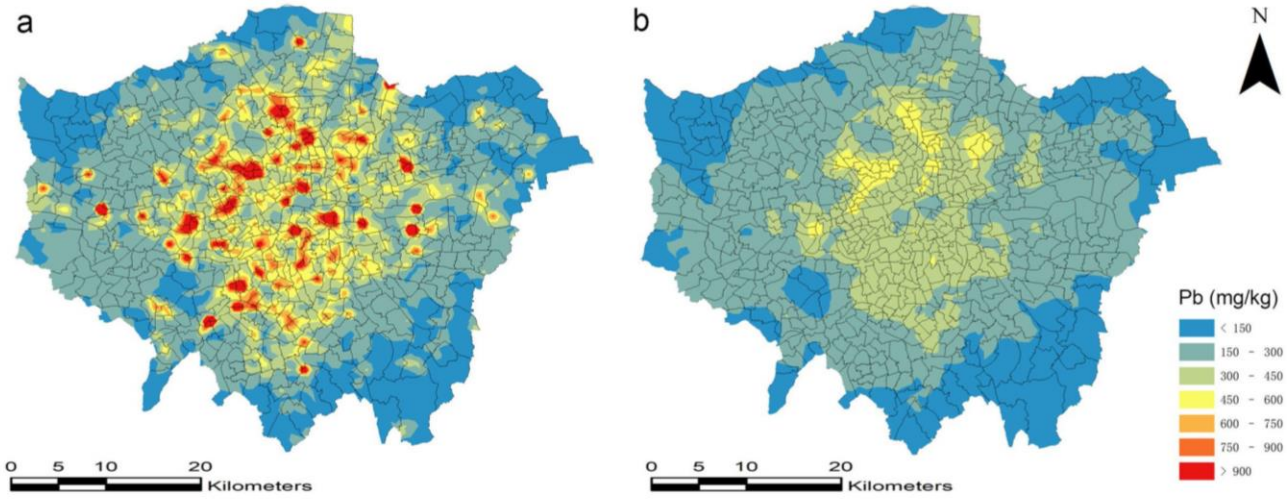


Fig. 4. Predicted topsoil Pb concentration (mg kg^{-1}) for the GLA area by using (a) IDW, (b) OK.

the smoothness of the estimates (Isaaks and Srivastava, 1990). In general, IDW is easier to use when comparing to OK. Typically, the choice of semivariance function is somewhat arbitrary and based on user experience. There is no doubt that OK has a powerful ability to predict the overall trend of soil distribution, while IDW has ability to reserve local maxima (hot spots) and local minima (cold spots) (Xie et al., 2011). On the other hand, the choice of interpolation parameters can have an impact on the prediction results. The parameters of IDW did not strongly influence the results, while the choice of variogram model would introduce substantial errors (Wong et al., 2004).

4. Conclusions

In this study, a variety of methodologies used in recent studies to link point and areal spatial data have been reviewed. Eight methods, (1) average, (2) median, (3) centroids IDW, (4) average block IDW, (5) median block IDW, (6) centroids OK, (7) average block OK and (8) median block OK, were compared for their performances of transforming point data of Pb concentration to areal data in GLA area. The median of the predicted values at the nodes of the grid within each ward was selected to describe Pb level for each ward, which would be used for further investigations of the association between Pb and socioeconomic or health status in the GLA area in our future work. This study has investigated and highlighted the point-to-area transformation problem while dealing with the compatibility of point data and areal data. Each method has its merits and shortcomings and researchers can adopt the 'optimal' one to suit their studies.

Acknowledgements

This study is sponsored by the Royal Society International Exchange Scheme of the UK (2015–2017). The soil data used in this study were provided by the British Geological Survey from the London Earth Project. Yuting Meng appreciated the financial support of China Scholarship Council (CSC).

References

Allen, M.A., Cave, M.R., Chenery, S.R.N., Gowing, C.J.B., Reeder, S., 2011. Sample preparation and inorganic analysis for urban geochemical survey soil and sediment samples. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 28–46.
Appleton, J.D., Cave, M.R., 2018. Variation in soil chemistry related to different classes and eras of urbanisation in the London area. Appl. Geochem. 90, 13–24.
Babak, O., Deutsch, C.V., 2009. Statistical approach to inverse distance interpolation. Stoch. Env. Res. Risk A. 23, 543–553.
Boutwell, B.B., Nelson, E.J., Emo, B., Vaughn, M.G., Schootman, M., Rosenfeld, R., Lewis, R., 2016. The intersection of aggregate-level lead exposure and crime. Environ. Res. 148, 79–85.
Cakmak, S., Hebbern, C., Pinault, L., Lavigne, E., Vanos, J., Crouse, D.L., Tjepkema, M., 2018. Associations between long-term PM2.5 and ozone exposure and mortality in the Canadian Census Health and Environment Cohort (CANCHEC), by spatial synoptic classification zone. Environ. Int. 111, 200–211.
Carlson, K., Neitzel, R., 2018. Hearing loss, lead (Pb) exposure, and noise: a sound approach to ototoxicity exploration. J. Toxicol. Env. Health Part B 21 (5), 335–355.
Cave, M., Wragg, J., Lister, R., 2018. The effect of lead in soil on crime deprivation in Derby, Leicester and Nottingham. Appl. Geochem. 88, 198–212.
Cesaroni, G., Badaloni, C., Gariazzo, C., Stafoggia, M., Sozzi, R., Davoli, M., Forastiere, F., 2013. Long-term exposure to urban air pollution and mortality in a cohort of more than a million adults in Rome. Environ. Health Persp. 121 (3), 324–331.
Chen, K., Liao, Q.L., Ma, Z.W., Jin, Y., Hua, M., Bi, J., Huang, L., 2015. Association of soil arsenic and nickel exposure with cancer mortality rates, a town-scale ecological study in Suzhou, China. Environ. Sci. Pollut. Res. 22, 5395–5404.
Cheng, Q., 2001. Multifractal and geostatistic methods for characterizing local structure and singularity properties of exploration geochemical anomalies. Earth Sci. 26, 161–164.
Cressie, N.A.C., 1993. Statistics for Spatial Data. Revised ed. Wiley, New York.
Davis, H.T., Aelion, C.M., Lawson, A.B., Cai, B., McDermott, S., 2014. Associations between land cover categories, soil concentrations of arsenic, lead and barium, and population race/ethnicity and socioeconomic status. Sci. Total Environ. 490, 1051–1056.
Fajčíková, K., Cvečková, V., Stewart, A., Rapant, S., 2014. Health risk estimates for groundwater and soil contamination in the Slovak Republic: a convenient tool for identification and mapping of risk areas. Environ. Geochem. Health 36, 973–986.
Farkhondeh, T., Samarghandian, S., Sadighara, P., 2015. Lead exposure and asthma: an overview of observational and experimental studies. Toxin Rev. 34 (1), 6–10.
Feigenbaum, J.J., Muller, C., 2016. Lead exposure and violent crime in the early twentieth century. Explor. Econ. Hist. 2016, 51–86.
Franke, R., Neilson, G., 1980. Smooth interpolation of large sets of scattered data. Int. J. Numer. Methods Eng. 15, 1691–1704. https://doi.org/10.1002/nme.1620151110.
Gharibvand, L., Shavlik, D., Ghamsary, M., Beeson, W.L., Soret, S., Knutsen, R., Knutsen, S.F., 2017. The association between ambient fine particulate air pollution and lung cancer incidence: results from the AHSMOG-2 study. Environ. Health Persp. 125 (3), 378–384.
Goldberg, M.S., Burnett, R.T., Stieb, D.M., Brophy, J.M., Daskalopoulou, S.S., Valois, M., Brook, J.R., 2013. Associations between ambient air pollution and daily mortality among elderly persons in Montreal, Quebec. Sci. Total Environ. 463–464, 931–942.
Gotway, C.A., Young, L.J., 2002. Combing incompatible spatial data. J. Am. Stat. Assoc. 97 (458), 632–648.
Great London Authority, 2016. 2016-based Trend Projection Results. GLA Intelligence Greater London Authority, City Hall Available online at:. https://data.london.gov.uk/dataset/projections-documentation, Accessed date: 26 June 2019.
Gump, B.B., Reihman, J., Stewart, P., Lonky, E., Granger, D.A., Matthews, K.A., 2009. Blood lead (Pb) levels: further evidence for an environmental mechanism explaining the association between socioeconomic status and psychophysiological dysregulation in children. Health Psychol. 28 (5), 614–620.
Hu, H., Dailey, A.B., Kan, H., Xu, X., 2013. The effect of atmospheric particulate matter on survival of breast cancer among US females. Breast Cancer Res. Treat. 139, 217–226.
Huang, J., Malone, B.P., Minasny, B., McBratney, A.B., Triantafilis, J., 2017. Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. Sci. Total Environ. 609, 621–632.
Huang, J., Li, G., Qian, X., Xu, G., Zhao, Y., Huang, J., Liu, Q., He, T., Guo, X., 2018. The burden of ischemic heart disease related to ambient air pollution exposure in a coastal city in South China. Environ. Res. 164, 255–261.
Huo, X., Li, H., Sun, D., Zhou, L., Li, B., 2012. Combining geostatistics with Moran's I analysis for mapping soil heavy metals in Beijing, China. Int. J. Environ. Res. Public Health 9 (3), 995–1017.
Isaaks, E.H., Srivastava, R.M., 1990. An Introduction to Applied Geostatistics. Oxford University Press, New York.
Johnson, C.C., 2011. Understanding the quality of chemical data from the urban environment – part 1: quality control procedures. In: Johnson, C., Demetriades, A., Locutura, J., Ottesen, R.T. (Eds.), Mapping the Chemical Environment of Urban Areas. Wiley-Blackwell, Oxford, pp. 61–76.
Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London.
Krige, D.G., 1966. Two dimensional weighted moving average trend surfaces for ore-evaluation. J. South. Afr. Inst. Min. Metall. 66, 13–38.
Lark, R.M., Scheib, C., 2013. Land use and lead content in the soils of London. Geoderma 209-210, 65–74.
Lee, C., Lee, Y., Lian, I., Su, C., 2016. Increased prevalence of esophageal cancer in areas with high levels of nickel in farm soils. J. Cancer 7 (12), 1724–1730.
Lee, H.A., Lee, W.K., Lim, D., Park, S.H., Baik, S.J., Kong, K.A., Jung-Choi, K., Park, H., 2015. Risks of lung cancer due to radon exposure among the regions of Korea. J. Korean Med. Sci. 30, 542–548.
Lin, W., Lin, Y., Wang, Y., Chang, T., Chiang, L., 2014. Assessing and mapping spatial associations among oral cancer mortality rates, concentrations of heavy metals in soil, and land use types based on multiple scale data. Int. J. Environ. Res. Public Health 11, 2148–2168. https://doi.org/10.3390/ijerph110202148.
Markus, J., McBratney, A.B., 2001. A review of the contamination of soil with lead II. Spatial distribution and risk assessment of soil lead. Environ. Int. 27, 399–411.
McDermott, S., Bao, W., Aelion, C.M., Cai, B., Lawson, A.B., 2014. Does the metal content in soil around a pregnant woman's home increase the risk of low birth weight for her infant? Environ. Geochem. Health 36, 1191–1197.
McKinley, J.M., Ofterdinger, U., Young, M., Barsby, A., Gavin, A., 2013. Investigating local relationships between trace elements in soils and cancer data. Spat. Stat 5, 25–41.
Mielke, H.W., Reagan, P.L., 1998. Soil Is an Important Pathway of Human Lead Exposure.
Mielke, H.W., Zahran, S., 2012. The urban rise and fall of air lead (Pb) and the latent surge and retreat of societal violence. Environ. Int. 43, 48–55 (Environ. Health Persp. 106, 217-229).
Mielke, H.W., Gonzales, C.R., Smith, M.K., Mielke, P.W., 2000. Quantities and associations of lead, zinc, cadmium, manganese, chromium, nickel, vanadium, and copper in fresh Mississippi delta alluvium and New Orleans alluvial soils. Sci. Total Environ. 246, 249–259.
Naicker, N., Richter, L., Mathee, A., Becker, P., Norris, S.A., 2012. Environmental lead exposure and socio-behavioural adjustment in the early teens: the birth to twenty cohort. Sci. Total Environ. 414, 120–125.
National Statistics data, 2015. https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london, Accessed date: August 2019.
Neamtiu, I., Bloom, M.S., Gati, G., Goessler, W., Surdu, S., Pop, C., Braeuer, S., Fitzgerald, E., Baciu, C., Lupsa, I.R., Anastasiu, D., Gurzau, E., 2015. Pregnant women in Timis County, Romania are exposed primarily to low-level (<10 μg/l) arsenic through residential drinking water consumption. Int. J. Hyg. Environ. Health 218, 371–379.
Núñez, O., Fernández-Navarro, P., Martín-Méndez, I., Bel-Lan, A., Locutura, J.F., López-Abente, G., 2016. Arsenic and chromium topsoil levels and cancer mortality in Spain. Environ. Sci. Pollut. Res. 23, 17664–17675.
Oliver, M.A., Webster, R., 2015. Basic Steps in Geostatistics: The Variogram and Kriging. 2211-808X vol. 1. Springer. https://doi.org/10.1007/978-3-319-15865-5.
Ouyang, W., Gao, B., Cheng, H., Hao, Z., Wu, N., 2018. Exposure inequality assessment for PM2.5 and the potential association with environmental health in Beijing. Sci. Total Environ. 635, 769–778.

Pearce, D.C., Dowling, K., Sim, M.R., 2012. Cancer incidence and soil arsenic exposure in a historical gold mining area in Victoria, Australia: a geospatial analysis. J. Expo. Sci. Environ. Epidemiol. 22, 248–257.
Pinichka, C., Makka, N., Sukkumnoed, D., Chariyalertsak, S., Inchai, P., Bundhamcharoen, K., 2017. Burden of disease attributed to ambient air pollution in Thailand: a GIS-based approach. PLoS One 12 (12), e0189909. https://doi.org/10.1371/journal.pone.0189909.
Poggio, L., Gimona, A., Spezia, L., Brewer, M.J., 2016. Bayesian spatial modelling of soil properties and their uncertainty: the example of soil organic matter in Scotland using R-INLA. Geoderma 277, 69–82.
Qin, J., Zhao, L., Chen, Y., Yang, K., Yang, Y., Chen, Z., Lu, H., 2015. Inter-comparison of spatial upscaling methods for evaluation of satellite-based soil moisture. J. Hydrol. 523, 170–178.
Rapant, S., Cvečková, V., Dietzová, Z., Fajčíková, K., Hiller, E., Finkelman, R.B., Škultétyová, S., 2014. The potential impact of geological environment on health status of residents of the Slovak Republic. Environ. Geochem. Health 36, 543–561.
Rapant, S., FajČíková, K., Cvečková, V., Ďurža, A., Stehlíková, B., Sedláková, D., Ženišová, Z., 2015. Chemical composition of groundwater and relative mortality for cardiovascular diseases in the Slovak Republic. Environ. Geochem. Health 37, 745–756.
Robertson, G., 1995. GS+: Geostatistics for the Environmental Sciences—GS + User's Guide Gamma Design Software (Plainwell).
Sahmel, J., Hsu, E.I., Avens, H.J., Beckett, E.M., Devlin, K.D., 2015. Estimation of hand-to-mouth transfer efficiency of lead. Ann. Occup. Hyg 59 (2), 210–220.
Schmidt, G., Pesch, R., Schröder, W., Conrad, A., Kolossa-Gehring, M., Feigenspan, S., Dobler, L., Wiesmüller, G.A., Birke, M., Utermann, J., 2011. The potential of spatial information in human biomonitoring by example of two German environmental epidemiology studies. Environ. Geochem. Health 33, 399–408.
Shaheen, A., Iqbal, J., 2018. Spatial distribution and mobility assessment of carcinogenic heavy metals in soil profiles using geostatistics and random forest, boruta algorithm. Sustainability 10, 799. https://doi.org/10.3390/su10030799.
Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 ACM National Conference, pp. 517–524 https://doi.org/10.1145/800186.810616.
Shi, X., Ayotte, J.D., Onda, A., Miller, S., Rees, J., Gilbert-Diamond, D., Onega, T., Gui, J., Karagas, M., Moeschler, J., 2015. Geospatial association between adverse birth outcomes and arsenic in groundwater in New Hampshire, USA. Environ. Geochem. Health 37, 333–351.
Son, J., Bell, M.L., Lee, J., 2010. Individual exposure to air pollution and lung function in Korea: spatial analysis using multiple exposure approaches. Environ. Res. 110, 739–749.
Sun, H., 2018a. Association of soil selenium, strontium, and magnesium concentrations with Parkinson's disease mortality rates in the USA. Environ. Geochem. Health 40, 349–357.
Sun, H., 2018b. Association of soil potassium and sodium concentrations with spatial disparities of prevalence and mortality rates of hypertensive diseases in the USA. Environ. Geochem. Health 40, 1513–1524.
Sung, C.Y., Park, C., 2018. The effect of site- and landscape-scale factors on lead contamination of leafy vegetables grown in urban gardens. Landsc. Urban Plan. 177, 38–46.
Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46 (Supplement), 234–240.
Tsai, K., Su, C., Chiang, C., Tseng, Y., Lian, I., 2017. Environmental heavy metal as a potential risk factor for the progression of oral potentially malignant disorders in central Taiwan. Cancer Epidemiol. 47, 118–124.
Veldhuizen, E.M., Stronks, K., Kunst, A.E., 2013. Assessing associations between socio-economic environment and self-reported health in Amsterdam using bespoke environments. PLoS One 8 (7), e68790. https://doi.org/10.1371/journal.pone.0068790.
VoPham, T., Bertrand, K.A., Tamimi, R.M., Laden, F., Hart, J.E., 2018. Ambient $PM_{2.5}$ air pollution exposure and hepatocellular carcinoma incidence in the United States. Cancer Causes Control 29, 563–572.
Wang, J., Wang, Y., Zeng, H., 2016. A geostatistical approach to the change-of-support problem and variable-support data fusion in spatial analysis. J. Geogr. Syst. 18, 45–66.
Wang, S., Huang, G.H., Lin, Q.G., Li, Z., Zhang, H., Fan, Y.R., 2014. Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. Int. J. Climatol. 34, 3745–3751.
Wang, Y., Li, M., Wan, X., Sun, Y., Cheng, K., Zhao, X., Zheng, Y., Yang, G., Wang, L., 2018. Spatiotemporal analysis of $PM_{2.5}$ and pancreatic cancer mortality in China. Environ. Res. 164, 132–139.
Wong, D.W., Yuan, L., Perlin, S.A., 2004. Comparison of spatial interpolation methods for the estimation of air quality data. J. Expo. Anal. Environ. Epidemiol. 14, 404–415.
Xie, Y., Chen, T., Lei, M., Yang, J., Guo, Q., Song, B., Zhou, X., 2011. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: accuracy and uncertainty analysis. Chemosphere 82, 468–476.
Yaghjyan, L., Arao, R., Brokamp, C., O'Meara, E.S., Sprague, B.L., Ghita, G., Ryan, P., 2017. Association between air pollution and mammographic breast density in the Breast Cancer Surveilance Consortium. Breast Cancer Res. 19 (36). https://doi.org/10.1186/s13058-017-0828-3.
Yao, L., Huo, Z., Feng, S., Mao, X., Kang, S., Chen, J., Xu, J., Steenhuis, T.S., 2014. Evaluation of spatial interpolation methods for groundwater level in an arid inland oasis, northwest China. Environ. Earth Sci. 71, 1911–1924.
Yeh, H., Hsu, S., Chang, Y., Chan, T., Tsou, H., Chang, Y., Chiang, P., 2017. Spatial analysis of ambient $PM_{2.5}$ exposure and bladder cancer mortality in Taiwan. Int. J. Environ. Res. Public Health 14, 508. https://doi.org/10.3390/ijerph14050508.
Young, L.J., Gotway, C.A., 2007. Linking spatial data from different sources: the effects of change of support. Stoch. Env. Res. Risk A. 21, 589–600.
Young, L.J., Gotway, C.A., Yang, J., Kearney, G., DuClos, C., 2009. Linking health and environmental data in geographical analysis: It's so much more than centroids. Spatial Spatio-temporal Epidemiol 1, 73–84.
Yuan, F., Li, X., Jowitt, S.M., Zhang, M., Jia, C., Bai, X., Zhou, T., 2012. Anomaly identification in soil geochemistry using multifractal interpolation: a case study using the distribution of Cu and Au in soils from the Tongling mining district, Yangtze metallogenic belt, Anhui province, China. J. Geochem. Explor. 116-117, 28–39.
Zahran, S., Mielke, H.W., Weiler, S., Gonzales, C.R., 2011. Nonlinear associations between blood lead in children, age of child, and quantity of soil lead in metropolitan New Orleans. Sci. Total Environ. 409, 1211–1218.
Zhang, C., Tang, Y., Xu, X., Kiely, G., 2011. Towards spatial geochemical modelling: use of geographically weighted regression for mapping soil organic carbon contents in Ireland. Appl. Geochem. 26, 1239–1248.
Zhao, H., Xia, B., Fan, C., Zhao, P., Shen, S., 2012. Human health risk from soil heavy metal contamination under different land uses near Dabaoshan Mine, southern China. Sci. Total Environ. 417-418, 45–54.
Zheng, X., Pang, L., Wu, J., Pei, L., Tan, L., Yang, C., Song, X., 2012. Contents of heavy metals in arable soils and birth defect risks in Shanxi, China: a small area level geographic study. Popul. Environ. 33, 259–268.

Chapter 5
Discussion

5.1 Overview of the Research Process

With the development of GIS technology, increasingly environmental and geophysical studies are applying it to investigate soil nutrient accumulation (Zhang et al., 2010; Qiu et al., 2011) and toxic heavy metal contamination (Severini et al., 2018; Kim and Choi, 2019). Eutrophication caused by over-enrichment of nutrients has been observed in London which has become an environmental concern. In addition, different metal elements display different tolerance to human disturbance which provide the information for determination of the contamination level. Therefore, the four papers applied GIS techniques to investigate the spatial distribution patterns and influencing factors of P, Al, Ca and Pb in topsoils of GLA at the regional scale and P in sediments of Lake Taihu at microscale. To achieve this objective, local Moran's I (section 4.1) was used to map the hotspots and high-low spatial outliers of P which successfully revealed the existence of over-enrichment of P and indicated geogenic and anthropogenic influences on the spatial distribution in an intensively urbanized area. Eutrophication is attributed not only by accumulated-P soils via leaching, but also by internal P of lake sediments. To further investigate the spatial structure of labile P in lake sediments, Moran' I and geostatistics were applied in section 4.2 to determine spatial variability of labile P quantitively. Local Moran's I and ANOVA in section 4.3 were also proven to be useful to identify different level of geogenic and anthropogenic influences on Al, Ca and Pb. Different interpolation techniques in GIS can be used to transfer point data to areal data and had respective advantages and disadvantages (section 4.4). The study of topsoils of London provides valuable information for policy makers. Meanwhile, the geostatistics was applied assess the spatial variability of labile P in sediments which can help to compare different DGT profiles from spatially quantitative perspective.

5.2 Contributions of Research

Discussion

The local Moran's I has been highlighted as a powerful tool for identification of the hot spots of P and enlightenment of exploring the influencing factors that control the spatial distribution pattern. There were some significant new findings. For example, the hot spot map provided remarkable evidence that tide affected the P concentrations in topsoils of lower Thames Estuary. This finding was supported by relative low concentration of P in lower estuary (Berbel et al., 2015; Gao et al., 2016) and high P in the coast (Fang, 2000). One other finding was that several hotspots scattered in the outskirts indicated the point sources from sewage treatment works, such as Beddington and Deephams. Neal et al. (2010) stated that the main concern of P contamination came from sewage within the UK.

The semivariogram and Moran's I are useful for quantifying spatial autocorrelation of labile P in binding gel at a submillimeter scale. The spatial analysis could add more valuable information than classic 1D descriptive statistics and provide new parameters for the DGT profile comparison. The quantitative parameters include nugget, sill, nugget/sill ratio, range, Moran's I. The sill indicates the variation level. The higher a sill value was, the larger its variation of labile P was. The range marks the maximum spatial autocorrelation distance. Locations further apart than this distance are spatially uncorrelated or independent. Summer was featured by small patches size and high heterogeneity which was suggested by short ranges. These parameters clearly showed the seasonal change of labile P in sediments, which was corresponding to the previous studies (Chen et al., 2016; Ding et al., 2016).

Local Moran's I was proven as a powerful tool for identifying the difference of spatial patterns among Al, Ca and Pb in an extensively urbanized city. The different proportions of the variance were distinguished by the results of the ANOVA and local Moran' I. The significant difference between built-up and non-built-up area supported the research of Appleton and Cave (2018). The distinction between within-buffer and outside-buffer was associated with historical leaded petrol from road traffic (Xue et al., 2017).

Last but not least, the incompatible scaling problem has been highlighted and the commonly used methods of integrating point and areal data have been reviewed. This review would provide more choices for environmental, biological scientists. Valuable comparison example and take-home messages have been offered for readers. The GIS methods for integrating point and areal data are a relatively popular and easy way to carry out transformation (Pinichka et al., 2017). In particular, the block method reduced estimation variance and provided more representative values of areal units (Young et al., 2009).

5.3 Advancement

In the field of soil contamination, there were three main advancements in this dissertation. Firstly, sewage treatment works played an important role in P contamination in London. Secondly, the accumulation of Ca in London city centre was associated with historical destruction of buildings. Thirdly, the Pb contamination was influenced by built-up activities and road traffic. These findings provided crucial information for soil management.

In the perspective of the utility of GIS-based techniques, there were two main advancements in this dissertation. One was the utility of geostatistics and GIS-based spatial analysis in profiles at submillimeter scale. The other was comparison GIS-based methods of transforming point data to areal data.

Reference

Appleton, J.D., Cave, M.R., 2018. Variation in soil chemistry related to different classes and eras of urbanisation in the London area. Appl. Geochem. 90, 13-24.

Berbel, G.B.B., Favaro, D.I.T., Braga, E.S., 2015. Impact of harbour, industry and sewage on the phosphorus geochemistry of a subtropical estuary in Brazil. Marine Poll. Bull. 93, 44-52.

Chen, M., Ding, S., Liu, L., Wang, Y., Xing, X., Wang, D., Gong, M., Zhang, C., 2016. Fine-scale bioturbation effects of tubificid worm (Limnodrilus hoffmeisteri) on the lability of phosphorus in sediments. Environ. Pollut. 219, 604-611.

Ding, S., Wang, Y., Wang, D., Li, Y., Gong, M., Zhang, C., 2016. In situ, high-resolution evidence for iron-coupled mobilization of phosphorus sediment. Sci. Rep. 6, 24341.

Fang, T.H., 2000. Partitioning and behaviour of different forms of phosphorus in the Tanshui Estuary and one of its tributaries, northern Taiwan. Estuar. Coast. Shelf Sci. 50, 689-701.

Gao, Z., Fang, H., Bai, J., Jia, J., Lu, Q., Wang, J., Chen, B., 2016. Spatial and seasonal distributions of soil phosphorus in a short-term flooding wetland of the Yellow River Estuary, China. Ecol. Inform. 31, 83-90.

Kim, S., Choi, Y., 2019. Mapping Heavy Metal Concentrations in Beach Sands Using GIS and Portable XRF Data. J. Mar. Sci. Eng. 7(2), 42. https://doi.org/10.3390/jmse7020042

Neal, C., Jarvie, H.P., Withers, P.J.A., Whitton B.A., Neal, M., 2010. The strategic significance of wastewater sources to pollutant phosphorus levels in English rivers and to environmental management for rural, agricultural and urban catchments. Sci. Total Environ. 408, 1485-1500.

Pinichka, C., Makka, N., Sukkumnoed, D., Chariyalertsak, S., Inchai, P., Bundhamcharoen, K., 2017. Burden of disease attributed to ambient air pollution in Thailand: A GIS-based approach. PLoS ONE 12(12), e0189909. https://doi.org/10.1371/journal.pone.0189909

Qiu, J., Li, H., Wang, L., Tang, H., Li, C., Ranst, E., 2011. GIS-model based estimation of nitrogen leaching from croplands of China. Nutr. Cycl. Agroecosyst. 90, 243–252.

Severini, M.D.F., Carbone, M.E., Villagran, D.M., Marcovecchio, J.E., 2018. Toxic metals in a highly urbanized industry-impacted estuary (Bahia Blanca Estuary, Argentina): spatio-temporal analysis based on GIS. Environ. Earth Sci. 77, 393. https://doi.org/10.1007/s12665-018-7565-5

Xue, S., Shi, L., Wu, C., Wu, H., Qin, Y., Pan, W., Hartley, W., Cui, M., 2017. Cadmium, lead, and arsenic contamination in paddy soils of a mining area and their exposure effects on human HEPG2 and keratinocyte cell-lines. Environ. Res. 156, 23-30.

Young, L.J., Gotway, C.A., Yang, J., Kearney, G., DuClos, C., 2009. Linking health and environmental data in geographical analysis: It's so much more than centroids. Spatial Spatio-temporal Epidemiol. 1, 73-84.

Zhang, Q., Yang, Z., Li, Y., Chen, D., Zhang, J., Chen, M., 2010. Spatial variability of soil nutrients and GIS-based nutrient management in Yongji County, China. Int. J. Geogr. Inf. Sci. 24(7), 965-981.

Chapter 6
Conclusion

6.1 Overview

This Ph.D. research is to apply GIS-based techniques to investigate the spatial variation of P in the topsoils of the GLA area and its association with the natural and anthropogenic factors. The submillimeter-scale spatial variation of labile P in sediments measured by the DGT has been analyzed using methods of geostatistics. The different level of natural and anthropogenic influences on Al, Ca and Pb has been identified based on GIS-based spatial analysis. The point-to-area transformation issue has been highlighted and a valuable example to compare the different GIS-based methods has been provided.

6.2 Main conclusions

Urban soil contamination remains a vital environmental and healthy concern because of large population. Meanwhile, the GIS-based techniques are powerful for handling with big data and extracting the characteristics and structure of spatial distribution. It is, therefore, of great importance to investigate the spatial distribution patterns of chemical elements. The identified sources that affected accumulation help the treatment and management of soil contamination. In addition, the further application in microscale profiles widened the utility of GIS-based techniques. Moreover, the comparison of different GIS-based methods for transforming point and areal data provided more alternatives for scientists with different purposes.

6.2.1 Spatial distribution patterns of P in topsoils of the GLA area and their natural and anthropogenic factors

1) Both natural and anthropogenic factors had impacts on the spatial variation of P in topsoils of the GLA area.

2) The spatial patterns of P were controlled by PMs, including alluvium and river terrace deposits.

3) The distribution of P along the lower Thames Estuary was affected by the tidal effect.

4) The accumulation of P was associated with population density and urbanization.

5) Sewage treatment works were the main point sources of P in urban soils. Fertilizer application also had an effect on the P concentration.

6) The hot spot map provided important information for soil management.

6.2.2 Submillimeter-scale heterogeneity of labile P in sediments characterized by the DGT and spatial analysis

1) The spatial distribution of labile P was visualized using GIS techniques. High values of labile P with strong spatial variation were displayed in summer. Low values of labile P with weak spatial variation were found in winter. Medium values of labile P with moderate spatial variation were observed in spring and autumn. The maximum labile P of winter existed at the bottom of the profile; the maximum labile P of spring and autumn existed at the middle of the map; and the maximum labile P of summer existed at the top of the profile.

2) The spatial variation of labile P was quantified using semivariogram and Moran's I. The labile P had significant submillimeter-scale spatial autocorrelation, regardless of seasonal change of spatial autocorrelation. The results suggested that the spatial analysis, semivariogram and Moran's I, could provide useful information about the 2D, submillimeter-scale spatial variation.

3) The spatial autocorrelation of labile P with strong spatial variations (small patches) in summer implied weak spatial dependence and complex spatial pattern, while the spatial autocorrelation of labile P with low spatial variations (large patches) in winter exhibited

strong spatial dependence and relatively uniform spatial pattern. The significant difference in summer and winter suggested the temperature strongly influenced the spatial patterns.

6.2.3 Identifying geogenic and anthropogenic controls on different spatial distribution patterns of Al, Ca and Pb in urban topsoil of the GLA area

1) Spatial analyses were useful in identifying spatial patterns of urban soil geochemistry.
2) Different levels of geogenic and anthropogenic controls found on different chemicals.
3) Al was mainly controlled by geology, elevated in clay PMs.
4) Pb was mostly under anthropogenic control, elevated in built-up area and near primary roads.
5) Ca was impacted by both geogenic and anthropogenic controls. The high Ca concentrations were in the White Chalk Subgroup of the southern GLA area as well as the historical deconstruction and reconstruction sites in the downtown area.

6.2.4 Comparison of methods for addressing the point-to-area data transformation to make data suitable for environmental, health and socio-economic studies

1) The point-to-area misaligned problem has been highlighted.
2) The methods for integrating point and areal data that were applied in recent literature have been reviewed.

3) Eight methods, (1) average, (2) median, (3) centroids IDW, (4) average block IDW, (5) median block IDW, (6) centroids OK, (7) average block OK and (8) median block OK, were compared for their performances of transforming point data of Pb concentration to areal data in the GLA area. The advantages and disadvantages of eight methods were pointed out.

4) The method (5), Median block IDW, was advised to calculate Pb concentration in each ward of the GLA area.

6.3 Recommendations

The results of this dissertation would provide several suggestions for carrying out soil management.

1) Sewage treatment works and golf courses should be paid more attention which have more risk of leaching P. Tree buffer could be applied around these places.

2) London Authority should notice the gardeners who plant vegetables particularly in urbanized backyards and allotments, to be cautious to Pb accumulation.

3) The management of lake sediments should be carried out in summer due to the high labile P release and its strong spatial variability.

6.4 Future research

According to the results obtained in this research, several recommendations for future research are put forward as follows:

1) Submillimeter-scale spatial variation of labile P in sediments measured by the DGT should be investigated at more sites, such as the locations with and without

macrozoobenthos bioturbation in order to compare the effects of macrozoobenthos bioturbation on P spatial distribution.

2) Submillimeter-scale spatial variation of labile P in sediments measured by the DGT could be investigated in a time series, and the temporal-spatial analysis could be available by GIS techniques.

3) The relationship between Pb level and socioeconomic or healthy status in the GLA area should be conducted, since the suitable method had been decided by the comparison of methods of converting point data into areal data.