Urban-rural gradients of POPs soil concentrations at a regional scale: quantification and prediction.

Shuai Song\textsuperscript{1}, Yonglong Lu\textsuperscript{1*}, Tieyu Wang\textsuperscript{1}, Sheng Zhang\textsuperscript{1}, Chao Su\textsuperscript{12}, Yajuan Shi\textsuperscript{1}, Qifeng Li\textsuperscript{12}, Zhaoyang Liu\textsuperscript{12}, Jing Meng\textsuperscript{12}

1. State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China. Corresponding author. E-mail: yllu@rcees.ac.cn (Yonglong Lu)

2. University of Chinese Academy of Sciences, Beijing 100049, China

*Corresponding author. E-mail: yllu@rcees.ac.cn

Abstracts

Quantitative study of urban-rural gradients for persistent organic pollutants (POPs) concentrations is extremely important for understanding POPs behavior and for ecological risk assessment and management. However, little has been studied about quantitative calculation for soil concentrations of POPs from urban to rural areas. In this paper, 153 soil samples were collected in the coastal region of Bohai Sea and Yellow Sea, and the distributions of polychlorinated biphenyls, Polyaromatic hydrocarbons, Perfluoroalkyl and polyfluoroalkyl substances were examined. Urban-rural gradient model (URGM) was derived by using atmospheric point source diffusion model combined with fugacity model to prove potential mathematical relationships.
among urban and rural soils, and the URGM was proved suitable for simulating urban-
rural soil concentrations through validation of measured data. Significantly linear
correlation was found between POPs amount in surface soil and city population,
between POP concentrations and the artificial surface area. Urban-rural POPs
concentrations were simulated by URGM model, and calibrated by urban population
and land-cover data. The results showed that it was suitable for simulating urban-rural
POPs concentrations at a regional scale. This study could provide a new method for
quantifying urban-rural gradients for POPs concentrations, and promote quantitative
research on coupling between land cover, socio-economic data and POPs
concentrations.

Graphic Abstract:

Key words: urban-rural gradient; POPs; land use; regional scale; spatial prediction;
coastal region

1. Introduction

The fate of persistent organic pollutants (POPs) has brought great concern from
considerable government officials and scientists, particularly when local emission of chemicals has resulted in dispersed contamination of large areas. Assessment of POPs state and impacts is very important for environmental management and human health risk management. However, the differences of chemical properties and regional emissions, heterogeneity of environmental media properties, spatially and temporally limited large-scale monitoring data have led to difficulties in the assessment of contaminants in regional scales. Based on measured data of POPs, quantitative research of distributed regularity in environmental media is extremely important for an understanding of what the fate of a chemical is at the regional scale.

Urban areas are generally regarded as major sources of some POPs to the surrounding regions. Pollutant concentrations in environmental media could be directly affected by huge differences of contaminant emission per unit area between urban and rural regions. For example, emissions of PAHs, PCBs and PBDEs in urban areas may be several times or even thousands of times higher than that in suburbs, which may result in POPs concentrations in urban soils several times or even tens of times higher than that in rural soils. However, little work was relevant to quantitative spatial distribution of POPs soil concentrations between urban and rural regions at a regional scale. As a follow-up of our previous study, this paper is intended to study differences in distribution and behavior of POPs between urban and rural areas.

At present, there are mainly two methods for describing POPs gradients from urban to rural regions. One is that a few representative studies about POPs transport simulation between urban and rural regions using fugacity model. For example,
Multimedia Urban Model (MUM) and Spatially Oriented MUM (SO-MUM) were used to simulate the fate of POPs from urban to suburban in Canada\textsuperscript{4, 6, 7}; the Berkeley-Trent-Urban-Rural Fate Model (BETR) was provided for Switzerland to investigate the origin of PCBs in air and to investigate their long-term fate and mass balance in the environment\textsuperscript{5}; the spatially resolved Berkeley-Trent-Urban-Rural Fate Model (BETR-UR) was designed by coupling land cover information to simulate the transport of POPs between urban and rural areas\textsuperscript{2}. The other method is that the analysis of numerous measured data through systematic description\textsuperscript{8, 9} or some classic atmospheric transport models\textsuperscript{10}. For example, ratios of urban to rural concentrations was used to assess urban-rural trends for PCBs\textsuperscript{11}, OCPs\textsuperscript{11} and PBDEs\textsuperscript{12}. Melymuk et al.\textsuperscript{10} used a radial dilution model to simulate the urban-rural gradients in air and proved the influence of the central city as a source of contaminants to the surrounding environment. However, most of them were not able to provide quantitative information about distribution or transport of POPs between urban and rural areas\textsuperscript{5}, and some research of POPs gradients were mainly focused on air concentrations\textsuperscript{10}. Soil as one of the important environmental media for POPs sink, quantitative study of spatial distribution of chemical concentrations in soils should also be paid more attention. According to our previous study, more than 95%, 91%, 53% of the total BaP\textsuperscript{13}, Phen\textsuperscript{2}, PFOS\textsuperscript{14} concentrations would be found in soils in Bohai coastal region which implied that soils would serve as the predominant sink of POPs. Similar results for PCBs and PBDEs were also found in Toronto\textsuperscript{6}, which could be attributed to the physical and chemical properties of different chemicals\textsuperscript{14}. Hence, a quantitative study
of urban-rural POPs gradients in soils should be conducted for better understanding of the POPs behavior and for ecological risk assessment and management of POPs at a regional scale.

In this paper, measured data of PAHs, PFASs and PCBs were used to study the concentration gradients of POPs between urban and rural areas. We will (1) first outline the potential factors affecting the spatial distribution and accumulation of POPs; (2) quantitatively describe what are the urban-rural gradients in the study area through theoretical analysis (formula deduction) and validation of measured data; (3) illustrate what extent of POPs in soils is attributed to land cover and population distribution; and (4) demonstrate that multi-driving factors and urban-rural gradients are applied to calibrate and predict the distribution of concentrations of PFASs, PAHs and PCBs in soils.

2. Material and Methods

2.1 Study area and sample collection

The study area consists of four provinces (Liaoning, Hebei, Shandong and Jiangsu) and one municipality (Tianjin) around the Bohai Sea and Yellow Sea. A total of 153 surface soils were collected from 22 coastal cities (Figure S1). The study area was divided into 50×50 km² grids, and 2~3 sites were included in each grid by considering site selection principle of representative land cover, spatial uniformity, regional environmental representation, and balanced distribution of sites. Top surface soil samples (0-10cm) were collected using a stainless steel trowel that had been rinsed with
methanol. Each sample was composed of five sub-samples that were collected from the center and four corners of an area of 100×100 m². Information about sampling sites including coordinates, land use and detailed descriptions were recorded in the field. Soil samples were transferred and stored in clean polypropylene (PP) zip lock bags. All samples were dried in air, homogenized with a porcelain mortar and pestle, sieved with a 2 mm mesh for PFASs and a 100 mesh for PAHs and PCBs, and stored in 250 mL PP bottles at room temperature until extraction.

Figure S1. Distribution of sampling sites and land cover information along Bohai Sea and Yellow Sea. (Supporting materials)
2.2 Analysis of samples and data

Analysis of PFASs, PAHs and PCBs, quality control and quality assurance were using the same methods as those described previously by Meng et al., Jiao et al. and Gao et al., respectively. Detail information are given in the SI.

Data analysis, including statistical analysis and t-test, was performed on SPSS 17.0. Linear correlation and urban-rural model simulation were calculated on EXCEL 2013, and maps of correlations were drawn on ORIGINLAB 2015. A confidence level of 95% was used for the statistical tests and correlation analysis. Land cover data for year 2010 at 30 m resolution was from GlobeLand30 (http://glc30.tianditu.com). Spatial data analysis and calculation, and spatial distribution map were all performed on ARCGIS 10.2. The flowchart of our study was shown as Figure 1.

Figure 1 Architecture map
3. Results and discussion

3.1 Contaminant concentrations and distribution in soils

A summary statistics of PCBs, PFASs and PAHs concentrations in soils is presented in Table S1 for comparison. Coefficients of variation for PCBs, PFASs and PAHs were 1.74, 0.80 and 2.11 (Table S1a), respectively, which indicated that spatial variability of the POPs levels in soils are large due to different natural and anthropogenic impacts, and the sampling sites from the highly contaminated areas resulted in a wide range of POPs levels in soils. Since our focus was put on the relationship between urban and rural soil levels but not on highly contaminated areas in this paper, the outliers were not taken into consideration (Table S1 b, c, d do not include 2 PCBs, 3 PFASs and 2 PAHs outliers). A reference value in Table S1 is the limitation of toxic substances in soils.

3.1.1 Polychlorinated biphenyls

The concentrations of PCBs ranged from 2.94 to 385.67 ng/g, with mean value of 19.89 ng/g (Table S1a). The mean levels of PCBs were larger than the global background concentration for 5.41 ng/g in soil, and also larger than mean soil concentration for 0.52ng/g across China where both urban and rural soils were sampled. Nevertheless, the values in the present study were comparable with other studies (average 1.07 - 35.5 ng/g) in East China, Northeast, South and North China. PCBs levels increased from barrenland and grassland to arable, forests and orchard, and those in urban soils are largest. PCBs levels in the urban soils (24.78 ng/g) were three times higher than that in grassland (8.55 ng/g) and barrenland soils.
(8.33 ng/g), which indicated that primary sources of PCBs in rural area might come from urban area through air transport. A high degree of land development with high population density in the study area implied that barrenland was relatively less affected by human activities, which may be the reason why the lowest PCBs levels were found in the barrenland soils. The mean PCBs concentration in the study area was close to the 20ng/g limitation, while the median was far lower than this value. The mean PCBs concentrations in soils from Tangshan, Qingdao, Tianjin, Rizhao, Cangzhou, Jinzhou and Zibo were larger than the reference value (Table S1b), which indicated that the soils in this region were subject to intensive domestic and industrial activities in and around the sampling sites.

Table S1 Descriptive statistics of PAHs, PCBs and PFASs in soils (Supporting materials)

<table>
<thead>
<tr>
<th>Values (ng/g)</th>
<th>PCBs</th>
<th>PFASs</th>
<th>PAHs</th>
<th>City</th>
<th>N*</th>
<th>PCBs</th>
<th>PFASs</th>
<th>PAHs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref. value</td>
<td>20.00**</td>
<td>-</td>
<td>1000.00***</td>
<td></td>
<td></td>
<td>20.00</td>
<td>-</td>
<td>1000.00</td>
</tr>
<tr>
<td>Mean</td>
<td>19.89</td>
<td>7.51</td>
<td>466.73</td>
<td>Tangshan</td>
<td>9</td>
<td>30.97</td>
<td>6.14</td>
<td>875.15</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.94</td>
<td>2.76</td>
<td>26.63</td>
<td>Dandong</td>
<td>9</td>
<td>14.98</td>
<td>6.42</td>
<td>818.70</td>
</tr>
<tr>
<td>Maximum</td>
<td>385.67</td>
<td>63.97</td>
<td>11463.52</td>
<td>Zibo</td>
<td>3</td>
<td>20.93</td>
<td>9.76</td>
<td>713.45</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>34.73</td>
<td>6.05</td>
<td>985.69</td>
<td>Tianjin</td>
<td>7</td>
<td>25.45</td>
<td>7.07</td>
<td>503.71</td>
</tr>
<tr>
<td>Variance</td>
<td>1206.29</td>
<td>36.57</td>
<td>971579.28</td>
<td>Dalian</td>
<td>11</td>
<td>18.36</td>
<td>6.75</td>
<td>498.07</td>
</tr>
<tr>
<td>Skewness</td>
<td>8.17</td>
<td>6.86</td>
<td>9.54</td>
<td>Binzhou</td>
<td>4</td>
<td>14.11</td>
<td>7.87</td>
<td>489.94</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>82.15</td>
<td>56.77</td>
<td>103.91</td>
<td>Nantong</td>
<td>6</td>
<td>10.93</td>
<td>8.27</td>
<td>478.31</td>
</tr>
<tr>
<td>Percentiles</td>
<td>10</td>
<td>5.11</td>
<td>4.28</td>
<td>Jinzhou</td>
<td>5</td>
<td>21.65</td>
<td>7.45</td>
<td>406.59</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>7.97</td>
<td>5.02</td>
<td>Qingdao</td>
<td>10</td>
<td>28.69</td>
<td>7.31</td>
<td>362.82</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>12.99</td>
<td>6.54</td>
<td>Yingkou</td>
<td>4</td>
<td>18.84</td>
<td>6.16</td>
<td>325.04</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>17.69</td>
<td>7.68</td>
<td>Lianyungang</td>
<td>6</td>
<td>9.04</td>
<td>6.95</td>
<td>323.35</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>36.89</td>
<td>10.25</td>
<td>825.74</td>
<td>Dongying</td>
<td>5</td>
<td>14.42</td>
<td>6.73</td>
</tr>
</tbody>
</table>

Distance from Edge of City:

<table>
<thead>
<tr>
<th></th>
<th>PCBs (151*)</th>
<th>PFASs (150*)</th>
<th>PAHs (151*)</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Huludao</td>
<td>Yancheng</td>
<td>Qinhuangdao</td>
<td>Cangzhou</td>
</tr>
<tr>
<td>0-2km</td>
<td>18</td>
<td>20.70</td>
<td>8.37</td>
<td>763.15</td>
</tr>
<tr>
<td>2-5km</td>
<td>17</td>
<td>26.64</td>
<td>7.26</td>
<td>412.09</td>
</tr>
<tr>
<td>5-7km</td>
<td>12</td>
<td>21.96</td>
<td>7.00</td>
<td>433.64</td>
</tr>
</tbody>
</table>
### Polyaromatic hydrocarbons

The mean concentration of PAHs in the study area was 466.73 ng/g that was significantly lower than the reference value (1000 ng/g). However, the PAHs concentrations of 12 soil samples are higher than the reference value, where 10 sampling points are located at or around the urban areas (<2 km), an outlier (11463.52 ng/g) at Cangzhou industrial zone, and an outlier (3803.45 ng/g) near the steel plant in...
Tangshan. According to guideline values for soil contamination \(^{26}\), this reference value is a threshold for human exposure risks. The mean PAHs concentrations in different cities showed that Tangshan, Zibo, Tianjin, Dandong and Dalian are relatively high, and those in Weihai and Rizhao are relatively low (Table S1b). The distribution of PAHs concentrations in soils in different cities are similar with the distribution of PAHs emissions reported by Shen et al. \(^{27}\). The mean value of PAH concentrations for urban soils (570.80 ng/g) is large, and this level is about half order of magnitude higher in forest (340.97 ng/g), arable (367.68 ng/g) and orchard soils (256.65 ng/g). The mean PAHs concentration in grassland soils is the largest (716.33 ng/g), however, these two sites may be more representative of the local soil concentrations. Higher levels of PAHs in grassland were also found in other research\(^{19}\), which indicated PAHs in the grassland may be easily accumulated in the grass soils. The mean PAHs concentration in barrenland soils is also the lowest, similar with that of PCBs, which showed less impact of human activities. However, all the different concentrations in different soils probably reflect a combination of the local heating sources, traffic as well as long-range transport. And also, we could speculate about the sources of elevated concentrations from rural area to urban area. PAHs in soils surrounding steel plants, power plants and industrial parks were found with high concentrations. Due to a number of heavily polluting factories have been moved from cities to their suburbs over the past decade, soil pollution in urban areas may be alleviated. However, since human population and heat of combustions in urban areas are relatively concentrated, PAHs concentrations in urban soils remain relatively higher, while rural areas may be mainly impacted by local straw
combustion and household coal burning in winter.

3.1.3 Perfluoroalkyl and polyfluoroalkyl substances

A distribution with lower level and narrower range was observed for PFASs compared with PAHs and PCBs. The concentrations of PFASs ranged from 2.76 to 63.97 ng/g, with a mean value of 7.51 ng/g (Table S1a). The levels of PFASs in the study area were significantly lower than those in Shanghai (mean value 10.4 ng/g\textsuperscript{28}), while significantly higher than those with mean value of 0.98 ng/g around Bohai Sea\textsuperscript{15}, which indicated that land use, population density, industrial emission, historical accumulation may lead to differences in the regional distribution. PFASs in urban soils had generally higher mean values than the other soils, while for arable, orchard, barrenland and forest soils the mean and median PFAS values stayed close. Similar distribution trend was also reported by Meng et al.\textsuperscript{15}. It reflects continuous accumulation of contaminants in soils in the study area has probably led to increasing concentration, and intensive industrial and domestic processes in urban areas might greatly affect the PFASs concentrations in the surrounding soils. The mean PFASs concentration in soils from Zibo was the largest (Table S1b), which may be associated with emissions of solid waste, sewage and waste gas from fluorine chemical related industries. Besides, the mean PFASs concentrations in soils from Nantong, Tianjin, Jinzhou, Qingdao and Weifang were relatively higher (from 7.07 to 8.27 ng/g), for urbanization relevant indicators such as industrialization, urban population, and GDP per capita are relatively larger. The level of urbanization and change of PFOS concentrations are positively correlated\textsuperscript{15} which may be the key factors for the
differences of spatial PFASs concentrations among the different cities.

3.2 Theoretical derivation of urban-rural gradient model in soils

Cities could be regarded as major sources of PAHs, PCBs and other POPs to the rural regions, which was reflected in huge differences in contaminant emissions between urban and rural areas. Lots of studies on POPs concentrations in different media were reported, but little was studied about POPs behavior between urban and rural areas at a regional scale. Quantitative studies on urban and rural concentrations of pollutants mainly focus on those in atmospheric media using ratios of urban to rural concentrations. The use of these ratios were proved to be suitable for simulation of concentrations in atmosphere. However, rather than simple ratios in air, in our study, the dependent distance functions adapted to soil concentration gradients between urban and rural areas were clearly demonstrated by using formula deduction and validation of a large number of measured data.

To better understand the mechanism of chemical transport between urban and suburban areas, we adopted the fugacity equations to describe the transport process from atmosphere to soil (Figure S2). In the model, the bulk Z values (mol/Pa/m3) which are specific to the capacity of a phase for a chemical, and D values (mol/Pa/h) which are used to quantify the intermedia transport and transformation processes of chemicals between the two compartments, were used to calculate POPs transport from air to soil (details of D and Z values were described by Mackay). We assume that the pollutants in soils are mainly from the interactions between the atmosphere and soil while neglect other processes such as sewage irrigation and solid
waste pollution. When the chemical concentrations in atmosphere and soil reach equilibrium in urban or rural area, the following relationship exists for soils:

\[ f_a(D_1 + D_2 + D_3 + D_4) = f_s(D_5 + D_6) \]  \hspace{1cm} (1)

Concentrations of chemicals in air and soil were as follows:

\[ C_a = f_a \times Z_a \]  \hspace{1cm} (2)

\[ C_s = f_s \times Z_s \]  \hspace{1cm} (3)

Where, \( f \) is fugacity (mol/Pa/m\(^3\)); \( C \) is concentration; subscript \( a \) is air; \( w, \) water; \( s, \) soil; \( Q, \) aerosol; \( D_1, D_2, D_3, D_4, D_5, D_6 \) are transportation process of pollutant diffusion from air to soil, wet deposition of gas, dry deposition, wet deposition, diffusion from soil to air, and degradation, respectively. According to the equations (1-3), relations of concentrations between soil and air when reaching a steady state could be expressed by the equation (4). The parameters in equation (4) could be calculated by the equations (4-14), using which we could analyze the potential factors that impact the chemical transport in the different areas or media 2.

\[ C_s = C_a \left( \frac{(D_1 + D_3)}{(D_1 + D_2 + D_3 + D_4)} \right) Z_s \]  \hspace{1cm} (4)

\[ Z_a = \frac{1}{(RT)} \]  \hspace{1cm} (5)

\[ Z_w = \frac{1}{H} = \frac{Z_a}{K_{aw}} \]  \hspace{1cm} (6)

\[ Z_Q = 10^{(log K_{oa} + log f_0 - 11.91)} \times Z_a \times \rho_Q \times 10^9 \]  \hspace{1cm} (7)

\[ Z_s = 0.41 \times K_{ow} \times Z_w \times f_{oc} \times \rho_p \]  \hspace{1cm} (8)

\[ D_1 = \frac{A}{(1/ (k_{va} \times Z_a) + 1/ k_{vw} \times Z_w)} \]  \hspace{1cm} (9)

\[ D_2 = A \times U_i \times Z_w \times (1 - f_w) \]  \hspace{1cm} (10)

\[ D_3 = A \times v_Q \times Z_Q \times U_p \times (1 - f_w) \]  \hspace{1cm} (11)
\[ D_4 = A \times S_r \times \nu_Q \times Z_Q \times U_r \times (1 - f_w) \]  

\[ D_5 = A \left/ \left( k_v a \times Z_a \right) + 1/ k_v w \times Z_w \right) = D_1 \]  

\[ D_6 = 0.693 \times V_s \times Z_s / \tau_s \]  

\( R \) is the gas constant (8.314 Pa m^3/mol K); \( T \), absolute temperature (K); \( f_{oc} \), the fraction for organic carbon; \( \rho_Q \), and \( \rho_p \), the density of aerosols and soil partials (kg/m^3); \( A \), the media interfacial area (m^2); \( K \), Partition coefficient; \( K_{Qa} \), aerosol/ air partition coefficient; \( K_{aw} \), aerosol/ water partition coefficient; \( K_{ow} \), octanol/water partition coefficient; \( k \), the mass transfer coefficient (MTC), (m/s); \( k_{va} \), the air side MTC (m/s); \( k_{vw} \), the water side MTC (m/s); \( \nu \), the volume fraction; \( \nu_Q \), aerosols volume fraction; \( \nu_a \), air volume fraction; \( \nu_w \), water volume fraction; \( \nu_s \), soil volume fraction; \( S_r \), the scavenging ratio, \( U_r \), the rain rate (m/h), \( U_p \), the dry deposition rate (m/h), \( f_w \), the canopy wet interception fraction, \( \tau_s \), half-life of concentrations in soils (h); \( V_s \), volume (m^3).

The gradient of chemical atmospheric concentrations caused by the point source could be generalized by using a radial dilution model:

\[ C_r = C_0 \times r^m \]  

where \( C_r \) is air concentration from a given radius \( r \) to the point source with concentration \( C_0 \), and similarly identification of point source origin of atmospheric emission could also use this type of model to calculate with spatially-distributed air concentration data. At a regional scale, urban areas are generally regarded as major sources of POPs to the neighboring rural areas, and also urban areas are relatively small compared with the rural area, so atmospheric emission from urban areas could be generalized as point source emission in a large area. Some studies showed that similar atmospheric
concentration gradients from urban to rural areas, e.g., ratios of urban-rural air concentrations \(^{30}\), urban-rural air comparisons \(^{33}\), and simulation results of spatially-distributed POPs concentrations in air \(^{4,6}\). Melymuk et al\(^{10}\) demonstrated that PAHs, PCBs and PBDEs concentrations in air could be simulated using logarithmic form of the equation (15) with measured spatially distributed air concentrations.

If \(C_o\) is the pollutant concentration in the urban center, the parameter \(m\) should have significantly spatial difference for different sizes of cities. In simple term, each city should calibrate the model parameters using the surrounding measured data, e.g. the data from urban center to rural area, which might need more measured data to support. If we assume all the chemical concentrations in the urban are the same, and then \(C_o\) is the mean value, \(r\) should be the distance from the city edge. We use the constant \(\varepsilon\) value to correct the distance for chemical concentration differences caused by different areas of cities. The equation (15) is corrected for air gradient in a regional scale using logarithmic transformation:

\[
\ln\left(\frac{C}{C_o}\right) = m \times \ln(\varepsilon r) \tag{16}
\]

\[
\ln\left(\frac{C}{C_o}\right) = m \times \ln(r) + m \times \ln(\varepsilon) \tag{17}
\]

where \(\varepsilon\), a constant close to 1, denotes the degree of influence of urban area differences on concentration gradient. \(\varepsilon\) value = 1 represents no significant effect of area differences in cities on chemical gradient; \(\varepsilon\) value < 1 means that highest concentration was in urban area, and it has some influences of area differences in cities on chemical gradient; \(\varepsilon\) value >1 represents chemical concentrations in suburban larger than in urban and rural regions probably due to industrial transfer from urban to
suburban areas. The parameter m is the slope of the linear regression between ln(Cr/C0) and ln(εr).

According to the equation (4, 15 and 16), the chemical concentrations in soil between urban and rural areas could be represented as follows:

\[
dC_u = \frac{(dD_1 + dD_2)Z_{ur}}{(dD_1 + dD_2 + dD_3 + dD_4)Z_{ur}} \times dC_r \times (εr)^m
\]

where dC and dD denote chemical soil concentration per unit area and transportation process per unit area among the urban-rural transects. The subscript ‘r’ of variable dC and dD represents rural region from a given radius r to the urban center, and urban region; and ‘u’ represents urban area. Usually, Z values were related to the capacity of a phase for a chemical, so Z values for urban or rural media could be considered as the same, and then we can get equation (19).

\[
\ln \left( \frac{dC_u}{dC_r} \right) = \ln \left( \frac{(dD_1 + dD_2)Z_{ur}}{(dD_1 + dD_2 + dD_3 + dD_4)Z_{ur}} \right) + m \times \ln(r) + m \times \ln(ε)
\]

However, Song et al.\(^2\) simulated the different transportation processes between urban and rural areas, which proved that parameters such as vegetation cover, soil properties, and atmospheric aerosol concentration significantly affect the chemical fate.

If we actually consider the most potential parameter differences\(^2\) resulting from the given radius r to the urban edge, parameters in equation (19) and equation (9-14) except for \(v_Q, f_w,\) and \(Z_s\) all could be summarized as constants. Then we get equation (20), in which a, b, c, e, f and g could be got by deducing equations 9-14 and 19. \(v_Q\) in lower air actually is the function of r, but the distributions of \(f_w\) and \(Z_s\) values depend on vegetation cover and soil property. So, in equation 20, if anyone wants to study the effect of land cover and soil property on chemical distribution in soil, this model is a
simple method. Value m could be got from the concentration gradient in air, and then the distributions of land cover and soil property would be considered to calculate all the soil concentrations. Of course, if we assume the $f_w$ and $Z_s$ were spatially homogeneous, and then $C_{rs}/C_{us}$ was the function of single variable $r$ (equation 20).

\[
\ln \left( \frac{dC_{rs}}{dC_{us}} \right) = \ln \left( a \frac{b + (c + ev_0 + f_w)(1 - f_w)}{b + gZ_s} dA\right) + m \times \ln(r) + m \times \ln(\varepsilon) \quad (20)
\]

\[
\ln \left( \frac{dC_{rs}}{dC_{us}} \right) = m \times \ln(r) + m \times \ln(\varepsilon) \quad (21)
\]

According to equation (19), if we assumed that the transportation process (D values) in urban areas had no difference with that in rural areas, then equation (21) and (17) were similar. As can be seen from equation (20), slope $m$ is unrelated to the transportation process such as atmospheric deposition, and in theory the slope $m$ for chemical concentration in soil and air is the same. We think it is important finding for quickly estimating atmospheric and soil concentrations with lower cost and less time due to mutual conversion relationship at a large scale. However, we were lack of the measured data to prove the reliability, and in this paper we want to confirm the large-scale adaptability of equation (21) and the possible affecting factors.

3.3 Urban-rural gradient and validation of model

In our study, PCB, PFAS, and PAH concentrations in the urban soils were significantly larger than those in rural areas (Table S1d), where the urban soil information of sampling sites was supported by the fieldwork including small down town, built-up area, and green space. Generally, the concentrations of POPs gradually decrease from urban to rural soils. Urban concentrations (0-2 km sites) were 3.5, 1.5 and 1.4 times higher than the rural concentrations (30-50 km sites) for $\Sigma$PAHs, $\Sigma$PCBs
and ΣPFASs, respectively (Table S1d). Mean PFAS and PAH concentrations in urban soils were the largest, which indicated that industrial and domestic emissions of PFASs and PAHs per unit area in cities were the largest. PAH concentrations in 5-7 km from the urban center were slightly greater than those in 2-5km, with relatively higher standard deviation of 432 ng/g. It implied relatively greater suburban industrial emissions in 5-7 km from the urban center, and the aggregation and non-uniformity of industrial distribution might be the main reason for the greater standard deviation of PAH concentration. Urban–rural variation in PFAS gradients showed no large PFAS emissions to air in suburbs, and PFASs in urban areas were the main sources at the regional scale. PCB concentrations in soils with 2-5km away from the urban center were the largest, and the possible reason was the industrial transfer from urban to suburban areas.

In order to study urban-rural gradient rule, we count the distance of each point from the nearest edge of city in ARCGIS software using 2010 land cover data, and then PCBs, PFASs and PAHs concentrations in different segmented distances were analyzed. Although the given heterogeneity of concentrations within an urban area was found in our study, we hardly explain a generalized characteristic of cities using only 6 urban sampling sites at a large region scale. Besides, urban areas accounted for a small proportion of the entire study area. Therefore, the urban-rural gradients were mainly concerned and each POP was considered the same in the urban area. Artificial urban surface and the surroundings with 2 km distance were all regarded as urban areas in this study due to industrial transfer from urban to suburban areas ².
Figure 2. A logarithmic relationship exists between distance from edge of urban areas and the other sampling soil concentrations. Slops were significant at p<0.05 based on a linear regression t-test.

Due to the elevated concentrations in the vicinity of the urban areas, we chose equation (21) to test the hypothesis that the urban areas were the main sources of chemical concentrations to the study region. We used the measured data to validate the urban-rural gradient model with a known source region (0-2km) and compare the degree to which that region influences concentrations of PFASs, PCBs and PAHs. According to the analysis results from more than 150 sampling sites, a significant relationship was found between ln(Crs/Cus) and ln(r) for PFASs, PCBs and PAHs (Figure 2), where r is the distance in km from the edge of urban areas and Cus is the average concentrations in 0-2km urban areas, which proves that urban-rural gradient model (equation 21) is applicable for simulating POPs concentrations in soils at the regional scale. PAHs decreased most rapidly with increasing distance from the urban area, dropping off by 71% at 30-50 km urban areas, while PCBs and PFASs decreased
by 34% and 28%, respectively. The different slopes of the regression for all compounds indicated that the degree of declining concentrations along the urban–rural gradient (Figure 2) may be affected by urban emissions, land cover, physicochemical properties, and background emissions in suburban and rural areas.

3.4 Relationship of soil concentration with population combined by land cover data

As mentioned above, soil concentration of PCBs, PFASs and PAHs generally decreased from urban to rural areas in the study region. Meanwhile, variability of POPs spatial distribution was very large (Table S1b) even at the same distance from the city center. There may be a lot of reasons resulting in such a huge difference in the concentrations of pollutants in the soil, for example, structure and intensity of POPs emissions, wind speed and direction, topography, population distribution, wastewater discharge, the amount of coal, land cover types, soil physical properties, soil organic carbon\textsuperscript{34}, POPs properties and so on. Among so many factors, population is regarded as one of the most important factors affecting POPs emission, e.g estimation of PAH emissions\textsuperscript{13, 27}, PFAS domestic emissions\textsuperscript{35, 36} and PCB emissions\textsuperscript{37, 38}. At the same time, many articles have demonstrated that the population and concentrations for POPs in different environmental media showed a positive correlation\textsuperscript{15, 39, 40}. Land cover is also one of the important factors that influence the distribution of POPs, and different types of land use have a very significant impact on the distribution of POPs\textsuperscript{41}. Therefore, in this study, we try to analyze the relationship between spatial concentration of
pollutants and anthropogenic activities by using representative parameters of population and land cover data.

The relationship of different city population and amount of POPs in surface soil was analyzed, which showed a significant linear correlation (raw data in Figure 3). The amount of POPs in each city was calculated according to mean concentration multiplied by the urban area, soil depth (10cm), and soil density that were assumed all the same in the study area. However, we found many outliers in Figure 3, such as the relatively high POPs amount in Dandong, Jinzhou, Dalian, Tangshan and Huludao, which may be caused by high emission per person, long-distance transportation, and long-term accumulation. We compared land use and sampling sites in these cities, and found that woodland and grassland areas in these cities were relatively large with a small degree of landscape fragmentation (Table S1 and Figure S1). According to the principles shown in Figure S2, chemical emissions and atmospheric deposition were more susceptible to the surrounding anthropogenic activities. Population densities around a large area of meadows and forests in the mountains were small, but those around arable land were higher because of widely distributed villages (Figure S1). If we did not consider the impact of wind speed, easy accumulation in meadows and woodlands with high soil organic matter, and long-distance transport, the arable land may accept more atmospheric deposition of POPs produced from the surrounding towns and villages. In this regard, we use the coupling land use data to make a conversion for the total amount of POPs, as shown in the following equation:

\[ TA = \sum_{k=0}^{n} C_{mean} \cdot \rho \cdot CA_k \cdot h + \sum_{l=0}^{n} C_{b,l} \cdot \rho \cdot CA_l \cdot h \]  

(22)
where, $T_A$, translated data of total amount (ng); $C_{\text{mean}}$, mean concentration for each city in urban and arable soils, ng/g; $C_{bg}$, background value of different chemicals, ng/g; $k$, land cover including artificial surface and arable land; $l$, land cover including grassland, woodland and barren land; $\rho$, soil density, 1.6g/cm$^3$; $CA$, area of land cover, cm$^2$; $h$, depth, 10cm.

Figure 3. The relations of POPs among soil, population and landscape patterns in different cities (a-c), and the relations between artificial surface area in 10km buffer area and concentration of POPs (d).

It could be seen from Figure 3, linear fitting between transformed data and
population in different cities got better, compared with the raw data. R² of PCBs, PFASs and PAHs were increased from 0.49, 0.40 and 0.28 to 0.71, 0.62 and 0.69, respectively. Compared the transformed data with raw data (Figure 3), all the chemicals in Dandong, Dalian and Tangshan were still above the trend line, indicating that the emissions per capita in those regions were higher than other regions.

In order to study whether the chemical concentrations and the surrounding land cover are interrelated, 2km, 5km, 10km and 20km buffer zones for each sampling site were made using ARCGIS software, and the area of artificial surface in each buffer zone was counted. We used artificial surface are to reflect cities and towns. Through statistical analysis of the relationship between soil concentration and the surrounding artificial surface area, we found that the artificial surface area within 10km buffer zone and pollutant concentrations had a significant linear relationship (Figure 3), and detailed statistical data were shown in Table S2. The linear correlation between artificial surface in the other buffer zones and soil concentration was not significant. The range of artificial surface area within 10km buffer zone was from 0 to 141 km². The soil concentrations of PAHs, PCBs and PFASs with 113 km² artificial surface surrounding the sampling site were 2.3, 2.3 and 1.5 times than those with 3 km² artificial surface. Compared with the linear relationship between PAHs and artificial surface, R² of PCB and PFAS were 0.88 and 0.87, respectively. This may be affected by the diversification of PAHs sources, capacity of long-distance transportation, concentrated discharge of industrial parks, and so on. The relationship between artificial surface area with 10km scale and the chemical concentration may be a good approach to quickly estimate the
spatial distribution of POPs at the regional scale.

3.5 Predicting the spatial distribution of POPs concentrations in soils

If PAH, PCB and PFAS concentrations in urban soil were known, Equation 21 could be used to calculate the concentration gradient in all suburbs. In this study, we first generated 5×5km grids (Total 8692 grids) using ARCGIS, and then the distance from the central point of each grid to the closest city edge was calculated. Not all of the 22 cities have monitoring data of POPs in urban soils, and a little measured data may not be representative for the POPs concentrations in the urban soils. So, the first step, the mean concentrations in this region were used as single parameter for the established urban–rural gradient model to calculate soil concentrations for 5×5km grid points. Spatial distribution of urban-rural POPs simulation results is shown in Figure 4a, and parameters for calculation were shown in SI-Step1. Second step, the mean concentration (parameter $C_{us}$) in each urban soil was calibrated according to the relationship between the population and the chemical amount in soils (Figure 3, a-c), while maintaining the total amount of pollutants unchanged in each city. The calibrated results for parameter $C_{us}$ were shown in Table S3. The spatial distribution of simulated concentration in urban and suburban areas for the second step was shown in Figure 4b. Finally, according to the artificial surface data in 10km buffer zones, chemical concentrations in each site were calibrated again, and then regional distribution map of contaminants was generated (Figure 4c). Calibrated process was shown in SI-Step3 and Table S4.

Predicted and measured data of PAHs, PCBs and PFASs from urban to rural areas
were shown in Figure 5, with the average relative errors of 12.84%, 17.77% and 2.81%, respectively. The statistical errors demonstrated that the precision of spatial simulation results for PAHs, PCBs and PFASs was relatively high using urban-rural gradient model with coupled population and land cover data. Compared this method with our previous BETR model\textsuperscript{13, 14}, it requires less parameters, easier and faster computation, but it can better characterize the gradients between rural and urban areas, with the accuracy of the simulation results comparable with that from BETR model\textsuperscript{2, 5, 14}. 
Figure 4 Predicted distribution of PAHs, PCBs and PFASs. (a) simulated results using urban-rural model at a regional scale. (b) simulated results calibrated by population combined with land cover data at a city scale. (c) simulated results calibrated by artificial surface at a 10km scale.
In this study, uncertainty of the simulation results may be affected by various assumptions when the prediction of soil POPs concentrations distribution was conducted, as some factors were not considered, such as wind speed and direction, centralized emission from industries, long-distance transmission, long-term accumulation, and seasonal change. These factors may result in prediction errors of concentrations in some sites. For example, the POPs concentrations in the surroundings of an industrial park with fluorine chemical plants, power plants and waste recycling plants may be higher than that in the other regions. In addition, there are higher organic matter contents (OMC) in the grassland and woodland soils, but whether the corresponding POPs concentrations are significantly related to OMC or not needs further verification.

Urban-rural gradient model was developed to simulate spatial distribution of POPs...
in soils combined with urban population and land-cover data in this paper. We found that the slope m for chemical concentrations in soil and air should be the same under certain assumptions, which may provide a method for predicting atmospheric – soil concentrations at a regional scale. Meanwhile, the model could be used either as a single method for prediction or being combined with multi-parameters such as distribution of emission for POPs, chemical properties, soil properties, industrial information, multimedia properties, GDP and other social parameters. For example, total organic carbon may be used to improve the accuracy of POPs concentrations in woodland or grassland with high humus. Simpler parameters, faster deployment capability, and high accuracy are the major features for this study, but more local and regional monitoring data are needed to simulate the mechanism of distribution, seasonal and annual changes of the urban-rural gradients with higher precision. This paper presented a new method coupled land cover, socio-economic data, and POPs concentration for quantitative study of POPs distribution at a regional scale

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 41501539, 414201040045), the International Scientific Cooperation Program (No. 2012DFA91150) and the National Major Science and Technology Projects for Water Pollution Control and Management (No. 2015ZX07203-005). The authors would like to thank the editors and reviewers for their valuable comments and suggestions.
Appendix A. Supplementary data

Reference


