Modelling Arsenic and Lead Surface Soil Concentrations using Land Use Regression

S. Deschenes[1] , E. Setton[2], P. A. Demers[3] and P. C. Keller[4]

[1] Department of Geography, University of Victoria, Victoria, Canada, steeve@carexcanada.ca
[2] Department of Geography, University of Victoria, Victoria, Canada, eleanor@carexcanada.ca
[3] Cancer Care Ontario, Toronto, Canada, paul.demers@cancercare.on
[4] Department of Geography, University of Victoria, Victoria, Canada, pkeller@uvic.ca

Abstract. Land Use Regression (LUR) models are increasingly used in environmental and exposure assessments to predict the concentration of contaminants in outdoor air. We explore the use of LUR as an alternative to more complex models to predict the concentration of metals in surface soil. Here, we used 55 soil samples of As and Pb collected in 1996 across British Columbia (BC), Canada by the Ministry of Environment. Predictor variables were derived for each sample site using Geographic Information System (GIS). For As ($R^2 = 0.44$), the resulting linear regression model includes the total length of roads (m) within 25 km, and bedrock geology. For the Pb model ($R^2 = 0.78$), the predictor variables are the total surface area of industrial land use (m^2) within 5 km , the emissions of Pb (t) within 10 and 25 km, and the presence of closed mines within 50 km. The study proposes that LUR can reasonably predict the concentrations of As and Pb in surface soil over large areas.

Key words: exposure, land use regression, geographic information systems, arsenic, lead

Introduction

Models that interpolate metal concentration in surface soils are often data and computer intensive, while demanding elaborate statistical analyses (Davis et al., 2009), and applied to small geographic areas. In air pollution modelling, simple linear regression, usually described as Land Use Regression (LUR), has been used to predict ambient concentrations using basic geographically derived independent variables, with comparable results to other more complex methods (Jerrett et al., 2005). More recently, Wu *et al.* (2010) also applied LUR approach to predict the concentration of Pb in soil within city.

This study uses the LUR method to predict concentrations of As and Pb in surface soil across BC, Canada. Model variables include natural and anthropogenic emission sources and controlling factors. Once developed, the LUR models can be used to predict concentration levels wherever the relevant geographical variables are available, and thereby derive a modelled concentration surface to aid in exposure assessment for health-related studies.

Methods

Data

Surface soil samples at 55 sites in BC were collected between 1994 and 1996 by the Ministry of Environment. The samples were analyzed by ICP-OES, after an *Aqua Regia* digestion. Multiple samples taken at the same location were averaged and values below detection limits were excluded. Table 1 shows the concentration value range, mean and median for both As and Pb.

Table 1. Statistic Descriptive (µg/g)

Metals	Minimum	Mean	Median	Maximum
Arsenic	0.25	5.50	6.88	25.70
Lead	1.63	9.90	45.66	621.20

At each sample location, the predictor variables (Table 2) were derived using a range of circular buffer sizes (100m to 50km). Other site specific variables were derived using the sample location value of the geographic variables (elevation, bedrock geology, etc.).

Modelling Approach

Table 2. Geographical variables

Name	Description	Sources and date
Roads (Highway, Major, Local, and Rails	Length (m) in buffer	DMTI Spatial (2007)
Land use (Residential, Commercial, and Industrial)	Area (m^2) in buffer	DMTI Spatial (2007)
Slope	Slope at sample site (%)	Natural Resources Canada (2002)
Elevation	Elevation at sample site (m)	Natural Resources Canada (2002)
Surficial Geology	Surficial geology at sample site	Geological Survey of Canada (1996)
Bedrock Geology	Bedrock geology at sample site	Geological Survey of Canada (1997)
Total Precipitation	Total annual precipitation at the closest station (mm)	Environment Canada (2006)
Population Density	Total population in buffer	Statistics Canada (2006)
Mines and Smelters	Total number of mines and smelters in buffer	BC MINFILE (2000)
National Pollution Release Inventory (NPRI)	Sum of total emission (t) of As and Pb to air, water, and soil in buffer	Environment Canada (1994 to 2010)

Concentrations were transformed using the natural logarithm because they were highly skewed to the right. Bivariate analysis identified the most correlated buffer size per category of explanatory variables with the concentration levels (log). The most correlated variables were included in the regression analysis for variable selection. The other site specific variables (elevation, bedrock geology, etc.) were also added to the first fitted models. Next, backward and stepwise selections were performed to identify the most significant variables. Then, variables that were collinear (variance inflator factor > 10), not significant, or had a relationship contrary to that expected were eliminated. The two

previous steps were repeated until the final models included only significant variables and all the regression assumptions were met, including normally distributed residuals with low spatial autocorrelation.

The sensitivity of the models was examined by bootstrap analysis to assess how the model parameters (coefficients and R^2) vary when the LUR models were sub-sampled. Random selections of the soil concentration data, with replacement, were repeated 10,000 times and the respective parameter results for 95% confidence interval were recorded. Statistical analyses were conducted using the open-source software R 2.14.1.

Results

The final LUR models explain the variation of the metals in the surface soil of As and Pb at 44% and 78%, respectively. Tables 3 and 4 summarize the model parameters. The significant predictors (the variance explained by the variable is in bracket) for As include the length of local roads (log) within 25km (28%), and the bedrock geology (9%); the parameters for Pb include industrial land use within 5km (10%), the industrial emissions of Pb within 25 km (log) (4%) and 10 km (7%), and the presence of closed mines within 50 km (12%).

Table 3. Regression Analysis Results for Arsenic

As Model	Value	SE	t	p
Intercept	-5.47857	1.36517	-4.013	0.0002
Length of Roads (m) within 25 km (log)	0.49839	0.09166	5.437	2.39e-6
Mafic intrusive rocks – dorite[a]	-1.36899	0.65123	-2.102	0.04143

[a]Other categorical variable values are not shown here

Table 4. Regression Analysis Results for Lead

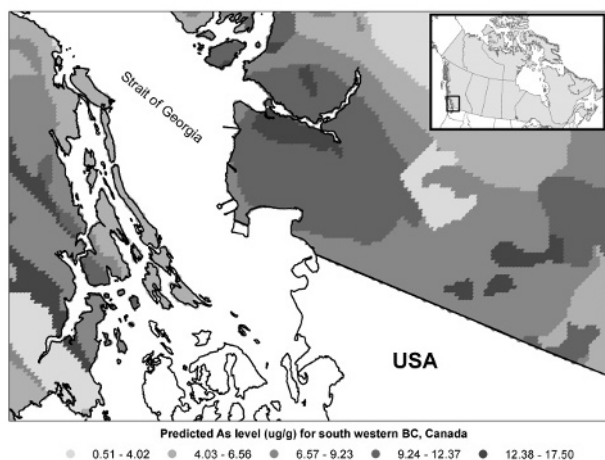| Pb Model	Value	SE	t	p
Intercept	1.82	1.35e-1	13.464	< 2e-16
Industrial land use within 5 km	1.49e-7	2.94e-8	5.069	5.62e-6
Industrial Emissions within 25 km (log)	9.76e-2	2.96e-2	3.303	0.00175
Industrial Emissions within 10 km	5.26e-5	1.21e-5	4.344	6.67e-5
Presence of Closed Mines within 50 km	1.199e-2	2.18e-3	5.501	1.23e-6

Table 5 shows the bootstrap 95% confidence intervals for the R^2 and the individual coefficient variables. The R^2 ranges from 0.3 to 0.9 with a mean of 0.63 for As and from 0.35 to 0.95 with a mean of 0.80 for Pb. For As, bedrock geology is sensitive to the resampling as it varies from a negative to a positive coefficient.

Table 5. Bootstrap Results for As and Pb

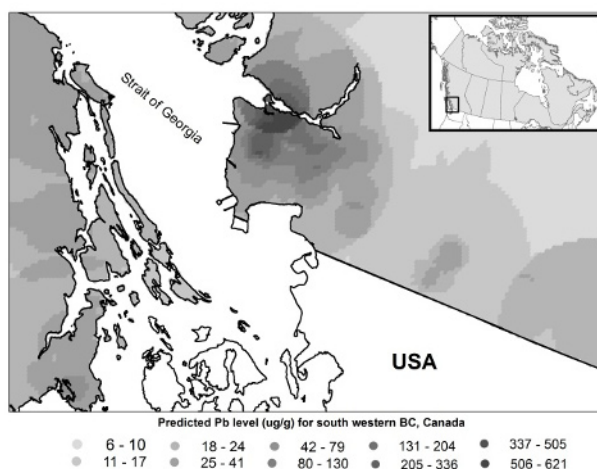| As Model	Model Value	95% CI[a]
R^2	0.44	0.30 to 0.89
Intercept	-5.479	-8.08 to -2.88
Length of Roads (m) within 25 km (log)	0.498	0.328 to 0.68
Mafic intrusive rocks – dorite[b]	-1.369	-1.521 to 0.447
Pb Model		
R^2	0.78	0.35 to 0.95
Intercept	1.82	1.58 to 2.14
Industrial land use (m²) within 5 km	1.49e-7	9.014e-8 to 2.119e-7
Industrial Emissions within 25 km (log)	9.76e-2	3.49e-2 to 1.51e-1
Industrial Emissions within 10 km	5.26e-5	2.56e-5 to 7.82e-5
Presence of Closed Mines within 50 km	1.199e-2	4.76e-3 to 1.58e-2

[a]Obtained after 10,000 iterations
[b]Only the significant factor is presented in this table

Fig. 1. Map of predicted As levels (μg/g)

Figures 1 and 2 show the predicted concentration surface for As and Pb in soil for south western BC. For the Pb model, the predicted values (27 points) above the

maximum level of the original concentration data were truncated to the highest concentration level of Pb because predicted values outside the concentration range might not be linear.

Fig. 2. Map of predicted Pb levels (μg/g)

Discussion

This study shows the potential of using LUR to predict the concentration of As and Pb in surface soil at regional levels. Overall model performance (As r2 = 0.44 and Pb r2 = 0.78) is similar to that obtained for air pollution models (Jerrett et al., 2005).

For the As model, bedrock geology reflects the natural source of the metalloid in soil, which agrees with other studies (Garcia-Sanchez and Alvarez-Ayuso, 2003; Grosz et al., 2004); however, the bedrock variable is qualitative and assumes continuous and homogenous As concentration in the bedrock unit, which may not be the case (Garcia-Sanchez and Alvarez-Ayuso, 2003). The road variable may be a surrogate for regional human activities related to the utilization of As-containing pesticide and wood preservative, combustion of fossil fuel (mainly coal plants), and other mining activities that are known to increase the concentration levels of As in the surrounding soils (Wang and Mulligan, 2006; Garcia-Sanchez and Alvarez-Ayuso, 2003). However, the relative higher mobility of As in soil may explain the lesser performance of the model compaired to Pb as past emissions may have leached down from the surface.

For Pb, the predominant predictors are the industrial land use surface area and the Pb emission point sources (NPRI). This finding agrees with other studies that demonstrated a strong influence of industrialization on the Pb surface soil concentrations (Murray et al., 2004; Salonen and Korkka-Niemi, 2007; Wu et al., 2010) and is reflected in the map (Figure 2). In fact, Wu *et al.* (2010) used a similar spatial approach than this study at city scale, and found strong relationships between Pb concentration levels in soil and proximity to roads and land use type. In opposition, this preliminary Pb model does not include any transportation related variables,

which is unexpected due to the legacy of Pb-gasoline emissions from vehicles, as reported by several studies (Adriano, 2001; Wu et al., 2010).

Unlike more complex dispersion models, the LUR approach does not typically consider factors controlling the deposition rate and direction from sources. Many studies reported that topography roughness, wind speed and direction influence atmospheric deposition, and thus the concentration of As or Pb in soil (Adriano, 2001; Fritsch et al., 2010).. The use of circular buffers to define geographic variables can result in patterns in the predicted surface that may not reflect actual distribution.

In this study, there is a discrepancy between the date of the data collection of the soil samples (1994 - 1996) and certain predictor variables (see Table 2). The predictor variables were used as indicators of human activities and we assume that the changes in the transportation network, population density, and land uses are minimal. For the NPRI, the emission levels are used as indicator of the location of the source. The emission data are reported from 1994, but we assume that the emission sources were active prior the beginning of the monitoring program and could be used as emission intensity indicators. Finally, the maps represent the situation at the time of the data collection (1994 - 1996). Any remediation or contamination done after the data collection is not accounted for in the present study.

Nonetheless, LUR soil models for As and Pb have a promising future as screening tools for remediation strategy and, more important, for human exposure assessments.

Acknowledgement

References

Adriano DC. Trace Elements in Terrestrial Environments: Biogeochemistry, Bioavailability, and Risks of Metals. Springer, 2001.

Davis HT, Aelion CM, McDermott S, Lawson AB. Identifying natural and anthropogenic sources of metals in urban and rural soils using GIS-based data, PCA, and spatial interpolation. Environmental Pollution 2009; 157: 2378-2385.

Fritsch C, Giraudoux P, Coeurdassier M, Douay F, Raoul F, Pruyot C, Waterlot C, de Vaufleury A, Scheifler R. Spatial distribution of metals in smelter-impacted soils of woody habitats: Influence of landscape and soil properties, and risk for wildlife. Chemosphere 2010; 81: 141-155.

Garcia-Sanchez A, Alvarez-Ayuso E. Arsenic in soils and waters and its relation to geology and mining activities (Salamanca Province, Spain). Journal of Geochemical Exploration 2003; 80: 69-79.

Grosz AE, Grossman JN, Garrett R, Friske P, Smith DB, Darnley AG, Vowinkel E. A preliminary geochemical map for arsenic in surficial materials of Canada and the United States. Applied Geochemistry 2004; 19: 257-260.

Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, Morrison J, Giovis C. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 2005; 15: 185-204.

Murray KS, Rogers DT, Kaufman MM. Heavy metals in an urban watershed in southeastern Michigan. J Environ Qual 2004; 33: 163-172.

Salonen VP, Korkka-Niemi K. Influence of parent sediments on the concentration of heavy metals in urban and suburban soils in Turku, Finland. Applied Geochemistry 2007; 22: 906-918.

Wu J, Edwards R, He X, Liu Z, Kleinman M. Spatial analysis of bioavailable soil lead concentrations in Los Angeles, California. Environmental Research 2010; 110: 309-317.