2009

Identifying Natural and Anthropogenic Sources of Metals in Urban and Rural Soils Using GIS-Based Data, PCA, and Spatial Interpolation

Harley T. Davis
C. Marjorie Aelion, *University of Massachusetts - Amherst*
Suzanne McDermott
Andrew B. Lawson

Published in final edited form as:
Environ Pollut. 2009 ; 157(8-9): 2378–2385. doi:10.1016/j.envpol.2009.03.021.

Identifying natural and anthropogenic sources of metals in urban and rural soils using GIS-based data, PCA, and spatial interpolation

Harley T. Davis[a], **C. Marjorie Aelion**[a,d,*], **Suzanne McDermott**[b], and **Andrew B. Lawson**[c,1]

a *University of South Carolina, Department of Environmental Health Sciences, 921 Assembly Street, Columbia, SC 29208, USA*

b *Department of Family and Preventive Medicine, 3209 Colonial Drive, Columbia, SC 29208, USA*

c *Department of Epidemiology and Biostatistics, 800 Sumter Street, Columbia, SC 29208, USA*

d *University of Massachusetts, Department of Public Health, 715 No. Pleasant Street, Amherst, MA 01003*

Abstract

Determining sources of neurotoxic metals in rural and urban soils is important for mitigating human exposure. Surface soil from four areas with significant clusters of mental retardation and developmental delay (MR/DD) in children, and one control site were analyzed for nine metals and characterized by soil type, climate, ecological region, land use and industrial facilities using readily-available GIS-based data. Kriging, principal component analysis (PCA) and cluster analysis (CA) were used to identify commonalities of metal distribution. Three MR/DD areas (one rural and two urban) had similar soil types and significantly higher soil metal concentrations. PCA and CA results suggested that Ba, Be and Mn were consistently from natural sources; Pb and Hg from anthropogenic sources; and As, Cr, Cu, and Ni from both sources. Arsenic had low commonality estimates, was highly associated with a third PCA factor, and had a complex distribution, complicating mitigation strategies to minimize concentrations and exposures.

Keywords

Principal component analysis; cluster analysis; rural and urban soils; GIS databases; neurotoxic metals; mental retardation and developmental delay

1. Introduction

Soils contaminated with toxic metals from point sources are potential exposure routes for surrounding populations (Carrizales et al., 2006; Hinwood et al., 2004; Pruvot et al., 2006). Surface soils are a relevant exposure route for a variety of metals (Caussy et al., 2003) such as arsenic (As) (Calderón et al., 2001; Díaz-Barriga et al., 1993), chromium (Cr) (Duckett, 1986), lead (Pb) (Calderón et al., 2001; Factor-Litvak et al., 1999), and mercury (Hg) (Debes

*Corresponding Author: (413) 545-2526 (phone); (413) 545-0501 (fax); E-mail: maelion@schoolph.umass.edu.
[1]Present address: Medical University of South Carolina, Department of Biostatistics, Bioinformatics, and Epidemiology, 135 Canon Street, Charleston, SC 29425, USA

et al., 2006), many of which have been associated with negative neurological impacts on humans. The relation between soil metal concentrations and residential dust metal concentrations also has been established (Hinwood et al., 2004; Hwang et al., 1997; Thornton et al., 1990; Wolz et al., 2003), regardless of whether the source is natural or anthropogenic, urban or rural.

It is often assumed that urban areas are more contaminated than rural areas due to the high number of these potential sources (Aelion et al., 2008; Harrison et al., 1981; Lau and Wong, 1982; Li et al., 2004; Madrid et al., 2002; Mielke et al., 1999; Surthland et al., 2000). Rural soils may also contain high concentrations of metals from natural geologic sources, pesticide application in high density agricultural areas (Micó et al., 2007; Rodríguez Martín et al., 2006; Wong et al., 2002; Yokel and Delistraty, 2003), localized industrial facilities, or atmospheric deposition (Nriagu and Pacyna, 1988). Understanding sources of metals in surface soils is necessary to implement mitigation strategies to reduce concentrations and limit human exposure.

The objectives of this study were to use publicly-available GIS-based data of eco-geologic, physico-chemical and industrial characteristics to determine commonalities of distribution, concentrations, and potential sources of nine metals in surface soils from four areas where a significantly high prevalence of mental retardation and developmental delay (MR/DD) was identified (two rural and two urban), and in a control rural reference area with no increased MR/DD prevalence compared to the state-wide average. The metals chosen were those with known environmental neurotoxicity such as As, Pb, Hg, and occupational toxicity (barium (Ba) (Jacobs et al., 2002), beryllium (Be) (Madl et al., 2007), Cr (von Berg and Liu, 1993), copper (Cu) (Stern et al., 2007), manganese (Mn) (Bouchard et al., 2008), and nickel (Ni) (Chashschin et al., 1994).

Map layers containing information on the natural geology of the area, soil type, climate division, ecosystem region, and anthropogenic characteristics of land use, presence of industrial facilities and Superfund sites were used to find associations with metal concentrations within each area. Kriging analysis of metal concentrations in each strip area was used to identify overlapping areas of higher metal concentrations. Because of the large numbers of sampling points within each area and the numerous metals measured at each point, principal component analysis (PCA) and cluster analysis were used reduce multidimensional data sets to lower dimensions (Barona and Romero, 1996) by identifying factors that contain the majority of the variance of the associated variables (Hill and Lewicki, 2005), and identify associations between metals in soils and possible commonalities of sources (Aelion et al., 2008; Li et al., 2004; Rodríguez Martín et al., 2006; Wong et al., 2002).

GIS-based data have been used in environmental soil (Jordan et al., 2007; Li et al., 2004) and ground water studies (Aelion and Conte, 2004), though most have not focused on the identification of possible contaminant sources and their potential associations to negative human health outcomes. Identifying potential commonalities in natural and anthropogenic metal concentrations and sources in surface soils in areas with known negative health outcomes may inform decisions on how to reduce contamination, and thereby minimize human exposure and protect populations at risk.

2. Materials and methods

Soil was collected from five areas. Strip 1 was a reference area with no increased prevalence of MR/DD compared to the state average, and Strips 2, 3, 4 and 5 had an increased prevalence of MR/DD identified using Bayesian local likelihood cluster modeling of maternal residence

during pregnancy and the MR/DD outcome of the child born to each mother (Zhen et al., 2008).

Strip 1 had the largest area (450 km^2) and Strip 5 had the smallest area (60 km^2); Strips 2, 3 and 4 were similar in size (150, 120 and 130 km^2, respectively). The numbers of cases (those diagnosed with MR/DD) and controls (those with no MR/DD diagnosis) varied between strips, but were relative to each strip area's population (i.e. Strip 3 had the fewest number of cases and controls, and Strips 4 and 5 had the greatest number). Each strip was defined as urban or rural based on the US Census Bureau. Urban areas are defined as those census blocks that have a population density of at least 500 people mi^{-2} (US Census, 2002). Rural areas are defined as those that are not located in an urban area. Strips 1, 2 and 3 were defined as rural areas and had small towns located within the strip areas. Strips 4 and 5 were defined as urban areas and had Metropolitan Statistical Area (MSA) populations of approximately 300,000 (Table 1).

Strip area latitude and longitude coordinates were mapped using ArcGIS® Version 9.2 software (ESRI, 1999–2005). A regular 120-node grid was laid out over each strip area; therefore, soil sampling did not coincide with the actual address location of each mother-child pair within the strip due to confidentiality concerns. Additional layers (roads, counties, etc.) were mapped in order to navigate to sampling locations. Samples were taken as close to grid nodes as possible, and permission was received for samples collected on private property.

Surface soil samples were collected at 5-cm depths after removing any plant debris, and placed in sterile Whirl-Pak® bags (Aelion and Davis, 2007; Aelion et al., 2008). For Strip 1, the number of samples analyzed for metals was 60, and for Strips 2, 3, 4 and 5 the number of samples analyzed for metals ranged from 114 to 120. Soil pH was measured in Strips 3, 4 and 5 with an Orion pH electrode (Thermo Fisher Scientific, Inc., Waltham, MA) after mixing 7 g of soil with 35 mL of a ~7 pH solution for 10 min. All samples were sent to an independent analytical laboratory (Pace Analytical) within a week of collection, acid digested and analyzed for As, Ba, Be, Cr, Cu, Pb, Mn and Ni (EPA Method 6010),(mg kg^{-1} dry weight (dw)) with inductively coupled plasma-atomic emission spectroscopy (ICP-AES). Mercury was analyzed using the cold vapor process to concentrate the Hg which reduced detection limits, and subsequently analyzed using ICP-AES (EPA Method 7471). Duplicate, control, spiked and blank samples were analyzed by Pace Analytical to assure appropriate QA/QC. The reporting limits were approximately 0.005 mg kg^{-1} dw for Hg and $\leq$0.5 mg kg^{-1} dw for all other metals. Any samples with measurements below the reporting limit were treated as non-detectable (ND) and were set to zero for all statistical analyses.

Several GIS-based data sets were used to describe the strip areas. 1) Soil type data from the National Resources Conservation Service (NRCS), United States Department of Agriculture (USDA) were collected using the online Web Soil Survey (Soil Survey Staff, 2008). 2) Climate division information (NCDC, 2005), average high and low temperatures, and rainfall for cities and towns (NCDC, 2008) located in strip areas were collected from the National Climactic Data Center (NCDC), which divides each state into climate regions based on average temperature, precipitation and drought conditions. 3) Ecological regions (ecoregions) were identified using the United Stated Geological Survey (USGS) map "Ecoregions of North Carolina and South Carolina" (Griffith et al., 2002), which categorizes areas by ecosystem as well as environmental resources. 4) Land use and land cover information (e.g. residential, forest, agriculture, etc.) was obtained from the US Environmental Protection Agency (EPA) (Anderson et al., 1976). Only the category which covered the majority of each strip area, specific to each data layer, was identified. For soils, however, the strip area was examined with the Web Soil Survey by dividing the strip into areas of <100 km^2 (if necessary) and recording the dominant soil types for each subsection; therefore, more than one type of soil was identified for each strip.

In addition to the layers described above, different types of industrial facilities and their locations were identified in the strip areas. The EPA's Toxic Release Inventory (TRI; EPA, 2008) for 2006 was used to identify facilities located in each strip area and their self-reported on-site releases of metals (Aelion et al., 2009). The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) database for South Carolina (SC) was used to locate potential and identified state and federal Superfund sites (SCDHEC, 2007). Limited chemical release data were available for CERCLA sites; therefore, no release data were reported.

All statistical analyses (except kriging of metal concentrations) were performed with SAS Version 9.1 software (SAS Institute, 2002–2003) and an α level <0.05 was used to determine significance. All metal concentrations were compared by strip using ANOVA. Soil metal concentrations (for each strip separately) were spatially interpolated using the ArcGIS® Version 9.2 software's kriging geostatistical analyst feature (excluding Be due to the high numbers of nondetectable (ND) samples in Strips 3, 4, and 5). The original kriging method with a standard search of 2–5 neighbors using a spherical model was used. For each metal, estimated metal concentrations were divided into four categories based on quartile data for all strips (25th, median, and 75th percentiles). Then, each layer was made semi-transparent so that layers could be overlaid to provide a spatial representation of overlaps of metal concentrations (darker areas). For PCA analyses of metal concentrations (excluding Be as above), only factors with eigen values >1 were used and promax, varimax and no factor rotations were compared; however, only varimax data were reported because this analysis "results in uncorrelated components", and factor loadings are equal to correlations between variables and factors (Hatcher, 1994). A cluster analysis (average linkage) was completed for each strip separately using all retained factors from the PCA, and a tree diagram was generated to identify metal clustering.

3. Results

GIS-based data for all strips are summarized in Table 1. The Web Soil Survey identified six different major soil types from west to east for the reference Strip 1. The Georgeville loam (GeC), Ailey sand (AeB), Nason loam (NaE), Norfolk loamy sand (NoB), and Chewlaca loam (Ch) all have the same soil pH (5.5). All except Ch are well drained and have at least 1.2 m to the water table; Ch is poorly drained and has a shallow ground-water table of only 0.15 m. Johnston muck (JoA), the most eastern soil type, has a lower soil pH (4.2) and is very poorly drained with about 0 m to the water table. The main differences in soils from Strip 1 are in parent materials (GeC and NaE: clayey from argillite; AeB and NoB: loamy marine deposits; and Ch and JoA: loamy alluvium) and 0–8 cm soil profiles (GeC, NaE, and Ch: loam; AeB: sand; NoB: loamy sand; and JoA: muck). In general, from west to east, pH and depth to the water table decreased for soils in Strip 1.

For Strip 2, the northern part of the strip had the majority soil type of Coronaca sandy clay loam (CoB), which is characterized by higher pH (6.5) as well as slightly different parent materials (clayey from hornblende gneiss with masses of gabbro-diorite), and 0–8 cm soil profile (sandy clay loam) than the other types found in the southern portions of Strip 2: Cecil sandy loam, 2–6 % slope (ClB2) and Cecily sandy loam, 6–10 % slope (ClC2). ClB2 and ClC2 have the same features (pH of 5.5, parent materials of clayey granite and gneiss, >2 m to water table and 0–8 cm soil profiles of sandy loam) with the only difference in slope.

Strip 3 had three major soil types: Leaf fine sandy loam (Lf), Chipley loamy sand dark surface (Cn), and Wehadkee-Chastain association (Wk). All Strip 3 soil types were more acidic (4.7–5.1) than other strips; all also drain very poorly and have ≤0.6 m to the water table. Strip 3 soils

have parent materials (Ls: clayey marine; Cn: sandy fluviomarine; Wk: loamy alluvium) and 0–8 cm soil profiles (Ls and Wk: fine sandy loam; Cn: sand) that were similar to Strip 1 soils.

Strip 4 major soil types were ClB2 and ClC2, similar to Strip 2. Strip 5 had two major soil types: both Cecil-Urban land complexes with different slopes (2–10 % for CuC and 10–25 % for CuE). These soil types, which are similar to ClB2 and ClC2, have a pH of 5.3, are well drained, have parent materials of clayey granite gneiss, >2 m to water table, and a 0–8 cm soil profile of sandy loam. The main difference in these soils as compared to those in Strip 1, 2, 3 and 4 is that urban land makes up ~50 % of the soil, which means human transported material is also considered a parent material of CuC and CuE soil types. The pH values of the soils reported in the soil survey were similar to those measured in our samples for Strips 3, 4 and 5. The average measured pH was 4.3, 5.8, and 5.5 for Strips 3, 4 and 5 (data not shown), respectively, compared well with the data-base values of the Web Soil Survey, of 4.7–5.1, 5.5, and 5.3, respectively.

The climate divisions were different for Strips 1, 2 and 3 and the same for Strips 4 and 5 (Table 1). However, all strips had similar temperature ranges (−1 to 32 °C) and average monthly precipitation (10–11 cm). Monthly precipitation was greatest in March for Strips 4 and 5, July for Strips 1 and 3, and January for Strip 2.

The major ecoregions identified for Strips 2, 4, and 5 were the same (Southern outer Piedmont), but both Strip 1 and 3 were located in different ecoregions (Sand hills and Carolina flatwoods, respectively; Table 1). The Southern outer Piedmont (Strips 4 and 5) is part of the Piedmont region, a transition zone between the mountains and the coastal plains of SC. The Southern outer Piedmont is less hilly and has less precipitation than the more northern, mountainous area of the state. The majority rock types are granite and gneiss with clayey soils. The Sand hills region (Strip 1) contains marine sands and clays that have been deposited over older rocks. The land is hilly and sands are the predominant soil type, allowing water to easily infiltrate to ground water. The Carolina flatwoods (Strip 3) are compromised of flat surfaces and poorly drained soils, which are loamy due to previous coverage by shallow ocean waters. The water table is also higher, leading to the presence of more freshwater swamps and marshes. In all ecoregions, specific soil types and parent materials corroborated the more detailed Web Soil Survey information.

The most prominent Anderson land cover for Strips 1, 2 and 3 was forests (mixed for Strips 1 and 2 and evergreen for Strip 3; Table 1). Lands categorized as forested must have a ≥10 % tree crown closure percentage and have not been developed for other land uses. Evergreen forests have predominately evergreen tree species, while mixed forests have both evergreen and deciduous species. Residential was the most common land use for the urban Strips 4 and 5. The residential categorization of Strips 4 and 5 is applied to urban and built-up land. Urban and built-up land requires much of the land to be covered with structures and is subdivided into more specific land uses, such as residential, commercial, transportation, etc.

Strip 4 had the most TRI facilities of all strips, and Strip 2, though rural, had a similar number of TRI facilities as Strip 4 (Table 1). Strip 1 TRI facilities reported the highest total on-site releases of Hg and Mn to air (2.3 and 16 kg, respectively; data not shown), and Cu and Pb to water of all strips (113 and 20 kg, respectively; data not shown); releases to air for Cu and Pb were also reported for Strip 1. Strip 2 TRI facilities reported the highest total on-site releases of Cu and Pb to air (30 and 95 kg, respectively; data not shown) for all strips; releases of Cr to air and water and Mn to air were also reported. The TRI facilities in Strip 4 released the highest amounts of Cr to air (20 kg), and Ni to air (4.5 kg) and water (2.3 kg) of all strips; releases of Cu and Pb to air were also reported (data not shown). Strip 3 had one TRI facility located within the strip area, and Strip 5, though urban, had no facilities reporting on-site

releases. Based on total on-site releases in SC for 2006, Cu and Pb had the highest on-site releases by TRI facilities (7000 and 4340 kg, respectively), followed by Mn, Cr, Ni, and Hg, ranging from 47 to 2930 kg (data not shown).

On-site release data were not available for the CERCLA sites, so only number of sites was compared between strips. Strip 4 also had the highest number of identified CERCLA sites (30) (Table 1); 15 of the sites were investigated by the state and one was characterized as a Brownfield site. Strips 1, 3, and 5 also had state-investigated sites (5 of 9 for Strip 1, 1 of 1 for Strip 3, and 11 of 13 for Strip 5). Both Strips 1 and 2 had one Brownfield site.

All metal concentrations were significantly different by strip (Table 2). Overall, Strips 2, 4, and 5 had higher average metal concentrations than Strips 1 and 3. Strip 5 concentrations of As, Cr, and Pb were significantly greater than those in Strips 1, 2 and 3 ($p<0.0001$). Strip 2, 4, and 5 concentrations of Ba, Cu, and Ni were significantly greater than Strips 1 and 3 ($p<0.0001$). Strip 2 Hg concentrations were significantly greater than Strips 3, 4 and 5 ($p<0.0001$). Beryllium concentrations were significantly greater for Strip 2 than Strips 3 and 5 ($p=0.0003$); however, these strips had high numbers of NDs (98 and 100 %, respectively). Strip 2 Mn concentrations were significantly greater than all other strips, and Strip 4 and 5 Mn concentrations were significantly greater than Strips 1 and 3 ($p<0.0001$).

Results of overlays of Kriged estimated metal concentration are shown in Figure 1. Darker colors indicate overlap of higher metal concentrations. For Strips 1 (Figure 1A) and 3 (Figure 1C), the majority of the strip areas are lightly colored, suggesting both low concentrations and little overlap of areas with high metal concentrations. In comparison, the majority of the areas in Strips 2, 4, and 5 (Figures 1B, 1D, and 1E, respectively) are dark in color, which suggests both higher concentrations of metals and spatial overlap of metals of high concentrations in these strips. This spatial representation corroborates results from ANOVA strip comparisons and GIS data (i.e. Strips 2, 4, and 5 are most similar and have highest metal concentrations).

The percents of variance explained by all retained factors from PCA (varimax rotation) are presented in Table 3, but no data are shown for specific factors of the PCA. In Strip 1, Factors 1 and 2 explained 55.6 and 21.2 % of the variance, respectively. For Strip 1, As, Ba, Be, Cr, Cu, Mn, and Ni were all highly correlated with Factor 1 (all correlations ≥ 0.71); Pb and Hg were highly correlated with Factor 2 (correlations of 0.8 and 0.88, respectively) Results of the cluster analysis for factors in Strip 1 (Figure 2A) showed similar results: Pb and Hg were clustered independently of other metals, while all other metals appear grouped; Ba, Be, Cu, Mn, and Ni appeared most tightly grouped of metals associated with Factor 1. In Strip 2, Factors 1 and 2 explained 27.6 and 21.9 % of the variance, respectively; a third factor was also retained in Strip 2 that explained 17.8 % of the variance (Table 3). Beryllium, Cr, Cu, and Ni were most associated with Factor 1 (all correlations ≥ 0.69) and Ba and Mn were highly correlated with Factor 2 (correlations ≥ 0.85). Arsenic, Pb, and Hg were most associated with the third factor (correlations of 0.85 and 0.81 for Pb and Hg, respectively); however, As had a correlation of only 0.32 with the third factor and was weakly correlated with all factors in general. The cluster analysis tree diagram for Strip 2 showed similar results (Figure 2B): As, Pb, and Hg were tightly grouped; Ba and Mn were closely grouped; and Be, Cr, Cu, and Ni were grouped. In Strip 3, Factors 1 and 2 explained 28 and 26.8 % of the variance, respectively (Table 3). Arsenic, Ba, Mn, and Hg were associated with Factor 1 (correlations ≥ 0.63) and Cr, Cu, Pb, and Ni were associated with Factor 2 (correlations ≥ 0.6). Cluster analysis results (Figure 2C) corroborated these groupings. Strip 4 had the most complex PCA results, with four factors retained explaining 23.7, 23.5, 15.2, and 14.4 % of the variance, respectively (Table 3). Chromium and Ni were associated with Factor 1, Ba and Mn with Factor 2, As and Pb with Factor 3, and Cu and Hg with Factor 4. All correlations were ≥ 0.6. Cluster analysis results for Strip 4 (Figure 2D) also showed these four metal groupings. In Strip 5, Factors 1 and 2 explained 30.1 and

20.3 % of the variance, respectively (Table 3). Barium, Cr, Mn, and Ni were most associated with Factor 1 while all other metals were associated with Factor 2. However, Cr had a low correlation with Factor 1 (0.37) while all others correlations for Factors 1 and 2 were ≥0.54 for Strip 5. The results of the cluster analysis for Strip 5 (Figure 2E) showed a slightly different pattern. Ba, Mn, and Ni were grouped, however Cr was grouped with As and Cu, and Pb and Hg were grouped separately. These groupings may be due to the fact that As, Cr, Cu, and Hg were not closely associated with either factor in Strip 5 (in comparison to most other strips). Of all strips, the most variance was explained by Factors 1 and 2 of Strip 1 (76.8 % total); however, all four factors of Strip 4 explained the same amount of variance. Strip 5 had the lowest total explained variance (50.3 %).

Communality is the proportion of variance of a particular item that is due to common factors (shared with other items). The proportion of variance that is unique to each item is, therefore, the respective item's total variance minus the communality (Statistica, 2008). Arsenic had the lowest communality estimate of all metals in Strips 1, 2, and 4, and Cr had the lowest communality estimates in Strips 3 and 5 (Table 4). Commonality estimates for Cr were also low in Strip 1, and commonality estimates were low for Hg in Strips 3 and 5. The highest communality estimates were for Ba and Mn in Strips 2, 3, and 5; for Be and Ni in Strip 1; and for Cr and Ni in Strip 4. Overall, Strip 1 had the highest and Strip 5 had the lowest communality estimates for all metals, which corroborates the total percent of variance explained by factors for each strip.

4. Discussion

Higher metal concentrations, based on both strip comparisons and spatial interpolation of metals by strip, were associated with Strips 2, 4 and 5, all of which had a higher MR/DD prevalence than the control strip. Strip 3 also had a cluster of MR/DD (although MR/DD case and control numbers were low in comparison the other strips), and it had low soil metal concentrations. Eco-geological characteristics of Strip 3 were unique; soils were the most acidic and had the least distance (≤0.6 m) to the water table. Unlike the well drained soils in all other strips, Strip 3 soils were poorly drained, therefore contact is maintained between ground water and soils. Since metal dissolution increases with soil acidity (Dijkstra et al., 2004), it is possible that the combination of acidic and poorly drained soils, a shallow water table, and greater rainfall during high temperature summer months may enhance leaching of metals to the ground water in Strip 3 as compared to the other strips. In this case there may be metal contamination, but not accumulation as might occur in the other strips. Wong et al. (2002) concluded that concentrations of the heavy metals Cu and Zn were greater in the crop soils than natural and rice paddy soils in South China because flooding in paddy soils enhanced dissolution of Mn oxides, resulting in leaching of Cd and Co.

Land cover and use were not dominant determinants for soil metal concentrations; both the urban residential Strips 4 and 5 and the rural forested Strip 2 had high metal concentrations. Therefore it should not be assumed that rural, non-highly agricultural areas are uncontaminated, which corroborates findings of Aelion et al. (2008). Both broad (e.g. ecoregion, which take into account both geological and ecological components) and specific (e.g. soil type and soil physical properties) natural characteristics appeared to be important determinants of soil metal concentrations than land cover and use. Ecoregion categorization, detailed soil characteristics, and the season of highest monthly rainfall (winter/early spring) were similar for Strips 2, 4, and 5, which had the highest metal concentrations. Parent soil material of granite and gneiss were more common in Strips 2, 4, and 5, and, based on Web Soil Survey information, specific soil types were the same for Strips 2 and 4 and similar to Strip 5. In contrast, marine or loamy deposit parent materials were more common in Strips 1 and 3, and the highest average monthly rainfall was in the summer.

Kriging of metal concentrations in Strips 1–5 showed that Strips 2, 4 and 5 had large areas over which high concentrations of metals overlapped, which was not the case for the control Strip 1 and the rural Strip 3. PCA analyses indicated, in general, that the metals Ba and Mn were consistently grouped and may represent naturally-occurring metals in all strips, which was corroborated by cluster analysis. No point sources of Ba were identified in any of our strips and statewide on-site releases of Ba in 2006 were only 125 kg. Two TRI air sources of Mn were identified in Strips 1 and 2, but it is unlikely that these limited releases of Mn had a significant impact on Mn soil concentrations. In contrast to our study, Loska and Wiechu a (2003) and Singh et al. (2005), using PCA, concluded that Mn was from a combination of both natural and anthropogenic sources in river surface sediments of urban, industrial areas.

PCA suggested that both Pb and Hg were derived primarily from anthropogenic sources corroborating results of Möller et al. (2005) and Rodríguez Martín et al. (2006). Except for Strip 3, Pb and Hg were not highly correlated or grouped in cluster analysis with naturally occurring Ba and Mn. Lead concentrations were higher in urban areas than rural areas suggesting that population-dependent characteristic such as vehicular traffic and land use may be important. TRI facilities in rural Strip 2 reported the highest Pb release to air. Mercury was equally distributed regardless of land use and industrial facility number suggesting ubiquitous long-range distribution from atmospheric releases as wet or dry deposition from manufacturing, coal-fired power plants and other combustion processes associated with industrial facilities (Nriagu and Pacyna, 1988). SC has 12 coal-fired power plants (American Coal Foundation, 2008), and sixty-two state water bodies (creeks, rivers, streams and lakes) spread over greater than half the area of SC have fishing advisories due to high Hg concentrations (SCDHEC, 2008), another indicator of wide-spread Hg occurrence.

PCA indicated that both natural and anthropogenic sources contributed to Cr, Cu, and Ni concentrations in soils and were strip dependent. In the rural Strip 1, these metals were grouped with naturally-occurring metals, like Ba and Mn. Factor 1 explained twice as much of the variance in the data for Strip 1 than for all other strips suggesting more homogeneity in this strip. Li et al. (2004) concluded that Cr was geochemically associated with the major elements such as Al, Fe and Mn, which may originate from the soil parental materials, and Rodríguez Martín et al. (2006) concluded that Cr and Ni were from natural sources in agricultural topsoils of Spain, controlled by the soil parent rock on a regional scale, and that Cu also was natural except for local anomalies, which were attributed to anthropogenic influence (Rodríguez Martín et al., 2006). Wong et al. (2002) suggested that Cu and Cr in crop, paddy and natural soils had common geochemical characteristics in the highly agricultural Pearl River Delta of South China.

Other research has concluded that Cu (Möller et al., 2005) and both Cu and Ni (Li et al., 2004) were from anthropogenic sources in urban areas. In the current study, for strips other than Strip 1, Cr, Cu and Ni (but predominantly Cr and Cu) appeared to be impacted by various sources. These metals were grouped with different metals in different strips. Both land use and point source releases from TRI facilities may impact Cr and Cu, which were generally higher in the urban areas and in areas with TRI facilities that emitted Cr and Cu to air (Cu in the rural Strip 2 and both Cr and Cu in Strip 4). PCA of Strips 2 and 5 indicated a combination of natural and anthropogenic sources for Cu (loading factors distributed between Factors 1 and 2), and natural sources for both Cr and Ni (loading factor distributed primarily in Factor 1). Strip 4 indicated a combination of natural and anthropogenic sources and additional variance for Cu and Cr (loading factors distributed between Factors 1, 2, and 3) and natural sources for Ni (loading factor distributed primarily in Factor 1). Strip 3 indicated a combination of natural and anthropogenic sources for Cu, Cr and Ni (loading factors distributed between Factors 1 and 2), but concentrations of all metals were low in this strip. These metals show high variability in concentrations and in potential sources.

Arsenic behaved differently than all other metals. It was associated with naturally occurring metals in Strips 1 and 3 (Factor 1), but appeared most closely grouped with Pb and Hg in Strip 2, Pb in Strip 4, and with Cu and Cr in Strip 5. In Strip 2, Factor 3 explained 13.8 % of the total variance for all metals, yet accounted for the greatest loading factor for As. Similarly, for Strip 4, Factor 3 explained 12 % of the variance for all metals, and accounted for the greatest loading factor (greater than the sum of loading factors for Factor 1 and Factor 2 combined) for As. Land use appeared to be important for As; Strips 4 and 5, with majority residential land use, had the highest As concentrations, although As was also present in the rural Strip 2. Arsenic is known to occur in pesticides such as lead arsenate (Yokel and Delistraty, 2003). However, As has been measured in soils throughout the state (Aelion et al., 2008). Arsenic was not released by any facilities in the strips and state-wide TRI releases of As were minimal for 2006 (8 lbs) (EPA, 2008). This suggests more complexity is associated with its distribution in surface soils and additional research is required to identify these sources.

5. Conclusions

Surface soil metal concentrations are a product of a variety of natural and anthropogenic sources. It should not be assumed, however, that rural, non-highly agricultural areas have low soil metal concentrations. Rural Strip 2 behaved more like urban Strips 4 and 5 and had greater metal concentrations than the other rural Strips 1 and 3. Some metals (Ba and Mn) were consistently associated with natural sources in both the rural and urban areas although concentrations varied by strip. Other metals (Pb and Hg) were consistently associated with anthropogenic inputs in both rural and urban areas. Finally, the third category of metals (e.g., As, Cu, Cr, and Ni), was associated with both natural and anthropogenic sources, which varied by location. In general ecological and soil characteristics, particularly soil type, appeared to contribute most to the higher concentrations of metals measured in Strips 2, 4 and 5, suggesting an important natural component to the measured metal concentrations. For some metals associated with anthropogenic inputs, land use and the density of industrial facilities appeared to play a role in the higher concentrations measured, e.g., TRI air releases of Cu and Cr. Arsenic had a complex distribution in rural and urban soils; it was associated with an additional factor and had lower commonality compared to other metals measured. In these situations where the actual source of a metal is not readily discerned, reducing metal concentrations and mitigating human exposure will be more difficult than for those metals with a unique source.

Acknowledgments

Funding for this research was provided by the National Institutes of Health, National Institute of Environmental Health Sciences, Grant No. R01 ES012895-01A1. We thank F. Nemeth, B. Bey, M. Engle, and S. Jayasinghe for help with soil sampling, and Y. Liu for help with statistical analyses.

References

Aelion CM, Conte BC. Susceptibility of residential wells to VOC and nitrate contamination. Environmental Science and Technology 2004;38:1648–1653. [PubMed: 15074670]

Aelion CM, Davis HT. Use of a general toxicity test to predict heavy metal concentrations in residential soils. Chemosphere 2007;67:1043–1049. [PubMed: 17140621]

Aelion CM, Davis HT, McDermott S, Lawson AB. Metal concentrations in rural topsoil in South Carolina: potential for human health impact. Science of the Total Environment 2008;402:149–156. [PubMed: 18538375]

Aelion CM, Davis HT, McDermott S, Lawson AB. Soil metal concentrations and toxicity: associations with distances to industrial facilities and implications for human health. Science of the Total Environment 2009;407:2216–2223. [PubMed: 19155049]

American Coal Foundation. 2008. Available online at http://www.teachcoal.org/index.html/ Accessed on [7/15/08]

Anderson, JR.; Hardy, EE.; Roach, JT.; Witmer, RE. U. S. Geological Survey Professional Paper 964. Washington, DC: U.S. Government Printing Office; 1976. A land use and land cover classification system for use with remote sensor data.

Barona A, Romero F. Distribution of metals in soils and relationships among fractions by principal component analysis. Soil Technology 1996;8:303–319.

Bouchard M, Mergler D, Baldwin ME, Panisset M. Manganese cumulative exposure and symptoms: a follow-up study of alloy workers. Neurotoxicology 2008;29:577–583. [PubMed: 18562007]

Calderón J, Navarro ME, Jimenez-Capdeville ME, Santos-Diaz MA, Golden A, Rodriguez-Leyva I, Borja-Aburto V, Díaz-Barriga F. Exposure to arsenic and lead and neuropsychological development in Mexican children. Environmental Research 2001;85:69–76. [PubMed: 11161656]

Carrizales L, Razo I, Téllez-Hernández JI, Torres-Nerio R, Torres A, Batres LE, Cubillas AC, Díaz-Barriga F. Exposure to arsenic and lead of children living near a copper-smelter in San Luis Potosi, Mexico: importance of soil contamination for exposure of children. Environmental Research 2006;101:1–10. [PubMed: 16171795]

Caussy D, Gochfeld M, Gurzau E, Neagu C, Ruedel H. Lessons from case studies of metals: investigating exposure, bioavailability, and risk. Ecotoxicology and Environmental Safety 2003;56:45–51. [PubMed: 12915139]

Chashschin VP, Artunina GP, Norseth T. Congenital defects, abortion and other health effects in nickel refinery workers. Science of the Total Environment 1994;148:287–291. [PubMed: 8029704]

Debes F, Budtz-Jørgensen E, Weihe P, White RF, Grandjean P. Impact of prenatal methylmercury exposure on neurobehavioral function at age 14 years. Neurotoxicology and Teratology 2006;28:536–547. [PubMed: 17067778]

Díaz-Barriga F, Angel Santos M, de Jesús Mejía J, Batres L, Yáñez L, Carrizales L, Vera E, del Razo LM, Cebrián ME. Arsenic and cadmium exposure in children living near a smelter complex in San Luis Potosí, Mexico. Environmental Research 1993;62:242–250. [PubMed: 8344231]

Dijkstra JJ, Meeussen JCL, Comans RNJ. Leaching of heavy metals from contaminated soils: an experimental and modeling study. Environmental Science and Technology 2004;38:4390–4395. [PubMed: 15382869]

Duckett S. Abnormal deposits of chromium in the pathological human brain. Journal of Neurology, Neurosurgery, and Psychiatry 1986;49:296–301.

Environmental Protection Agency (EPA). EPA Toxic Release Inventory (TRI) Explorer Version 4.7. 2008. Available online at http://www.epa.gov/triexplorer/ Accessed online [7/1/08]

ESRI. ArcGIS Software Version 9.1. Redlands, CA: 1999–2005.

Factor-Litvak P, Wasserman G, Kline JK, Graziano J. The Yugoslavia prospective study of environmental lead exposure. Environmental Health Perspectives 1999;107:9–15. [PubMed: 9872712]

Griffith, GE.; Omernik, JM.; Comstock, JA.; Schafale, MP.; McNab, WH.; Lenat, DR.; MacPherson, TF.; Glover, JB.; Shelburne, VB. Ecoregions of North Carolina and South Carolina, (color poster with map, descriptive text, summary tables, and photographs). Reston, Virginia: U.S. Geological Survey; 2002. (map scale 1:1,500,000)

Harrison RM, Laxen DPH, Wilson SJ. Chemical associations of lead, cadmium, copper, and zinc in street dusts and roadside soils. Environmental Science and Technology 1981;15:1378–1383.

Hatcher, L. A step-by-step approach to using SAS™ for factor analysis and structural equation modeling. SAS Institute, Inc.; Cary, NC, USA: 1994. p. 586

Hill, T.; Lewicki, P., editors. Statistics: Methods and Applications. StatSoft Inc.; Tulsa, Oklahoma, USA: 2005.

Hinwood AL, Sim MR, Jolley D, de Klerk N, Bastone EB, Gerostamoulos J, Drummer OH. Exposure to inorganic arsenic in soil increases urinary inorganic arsenic concentrations of residents living in old mining areas. Environmental Geochemistry and Health 2004;26:27–36. [PubMed: 15214611]

Hwang YH, Bornschein RL, Grote J, Menrath W, Roda S. Environmental arsenic exposure of children around a former copper smelter site. Environmental Research 1997;72:72–81. [PubMed: 9012374]

Jacobs IA, Taddeo J, Kelly K, Valenziano C. Poisoning as a result of barium styphnate explosion. American Journal of Industrial Medicine 2002;41:285–288. [PubMed: 11920972]

Jordan C, Zhang C, Higgins A. Using GIS and statistics to study influences of geology on probability features of surface soil geochemistry in Northern Ireland. Journal of Geochemical Exploration 2007;93:135–152.

Lau WM, Wong HM. An ecological survey of lead contents in roadside dusts and soils in Hong Kong. Environmental Research 1982;28:39–54. [PubMed: 7106075]

Li X, Lee S, Wong S, Shi W, Thornton I. The study of metal contamination in urban soils of Hong Kong using a GIS-based approach. Environmental Pollution 2004;129:113–124. [PubMed: 14749075]

Loska K, Wiechula D. Application of principal component analysis for the estimation of source of heavy metal contamination in surface sediments from the Rybnik Reservoir. Chemosphere 2003;51:723–733. [PubMed: 12668031]

Madl AK, Brown JL, Kolanz ME, Kent MS. Exposure-response analysis for beryllium sensitization and chronic beryllium disease among workers in a beryllium metal machining plant. Journal of Occupational and Environmental Hygiene 2007;4:448–466. [PubMed: 17474035]

Madrid L, Díaz-Barrientos E, Madrid F. Distribution of heavy metal contents of urban soils in parks of Seville. Chemosphere 2002;49:1301–1308. [PubMed: 12489727]

Micó C, Peris M, Recatalá L, Sánchez J. Baseline values for heavy metals in agricultural soils in an European Mediterranean region. Science of the Total Environment 2007;378:13–17. [PubMed: 17306340]

Mielke HW, Gonzales CR, Smith MK, Mielke PW. The urban environment and children's health: soils as an integrator of lead, zinc, and cadmium in New Orleans, Louisiana, USA. Environmental Research 1999;81:117–129. [PubMed: 10433843]

Möller A, Müller HW, Abdullah A, Abdelgawad G, Utermann J. Urban soil pollution in Damascus, Syria: concentrations and patterns of heavy metals in the soils of the Damascus Ghouta. Geoderma 2005;124:63–71.

National Climactic Data Center (NCDC), National Oceanic and Atmospheric Administration (NOAA) Satellite and Information Services. Climatological Data Annual Summary; South Carolina. 2005.

National Climactic Data Center (NCDC), NOAA. 2008. Available online at http://www.ncdc.noaa.gov/ accessed on [7/1/08]

Nriagu JO, Pacyna JM. Quantitative assessment of worldwide contamination of air, water and soils by trace metals. Nature 1988;333:134–139. [PubMed: 3285219]

Pruvot C, Douay F, Hervé F, Waterlot C. Heavy metals in soil, crops and grass as a source of human exposure in the former mining areas. Journal of Soils and Sediments 2006;6:215–220.

Rodríguez Martín JA, López Arias M, Grau Corbí JM. Heavy metal contents in agricultural topsoils in the Ebro basin (Spain). Application of the multivariate geostatistical methods to study spatial variations. Environmental Pollution 2006;144:1001–1012. [PubMed: 16580763]

SAS Institute. SAS Software Version 9.1. Cary, NC: 2002–2003.

Singh KP, Malik A, Sinha S, Singh VK, Murthy RC. Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal component analysis. Water, Air, and Soil Pollution 2005;166:321–341.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. 2008. Available online at http://websoilsurvey.nrcs.usda.gov/ accessed [07/01/2008]

South Carolina Department of Health and Environmental Control (SCDHEC). Database of State and Federal CERCLA sites, State Priority List (SPL) sites and sites under current investigation. 2007. Available online at http://www.scdhec.gov/gis/gistopic.aspx/ accessed [5/1/08]

South Carolina Department of Health and Environmental Control (SCDHEC). South Carolina Fish Consumption Advisories. 2008. Available online at http://www.scdhec.gov/environment/water/fish/index.htm/ accessed on [7/5/08]

Statistica. Principal Components and Factor Analysis. StatSoft, Inc; 2008. Available online at http://www.statsoft.com/textbook/stfacan.html/ accessed [08/29/2008]

Stern BR, Solioz M, Krewski D, Aggett P, Aw TC, Baker S, Crump K, Dourson M, Haber L, Hertzberg R, Keen C, Meek B, Rudenko L, Schoeny R, Slob W, Starr T. Copper and human health: biochemistry, genetics, and strategies for modeling dose-response relationships. Journal of Toxicology and Environmental Health, Part B 2007;10:157–222.

Surthland RA, Tolosa CA, Tack FMG, Verloo MG. Characterization of selected element concentrations and enrichment ratios in background and anthropogenically impacted roadside areas. Archives of Environmental Contamination and Toxicology 2000;38:428–438. [PubMed: 10787093]

Thornton I, Davies DJA, Watt JM, Quinn MJ. Lead exposure in young children from dust and soil in the United Kingdom. Environmental Health Perspectives 1990;89:55–60. [PubMed: 2088756]

United States Census Bureau. 2002. http://www.census.gov/geo/www/ua/ua_2k.html Census 2000 Urban and Rural Classification

von Berg R, Liu D. Chromium and hexavalent chromium. Journal of Applied Toxicology 1993;13:225–230. [PubMed: 8326094]

Wolz S, Fenske RA, Simcox NJ, Palcisko G, Kissel JC. Residential arsenic and lead levels in an agricultural community with a history of lead arsenate use. Environmental Research 2003;93:293–300. [PubMed: 14615240]

Wong SC, Li XD, Zhang G, Qi SH, Min YS. Heavy metals in agricultural soils of the Pearl River Delta, South China. Environmental Pollution 2002;119:33–44. [PubMed: 12125727]

Yokel J, Delistraty DA. Arsenic, lead, and other trace elements in soils contaminated with pesticide residues at the Hanford Site (USA). Environmental Toxicology 2003;18:104–114. [PubMed: 12635098]

Zhen H, Lawson AB, McDermott S, Pande Lamichhane A, Aelion CM. A spatial analysis of mental retardation of unknown cause and maternal residence during pregnancy. Geospatial Health 2008;2:173–182. [PubMed: 18686266]

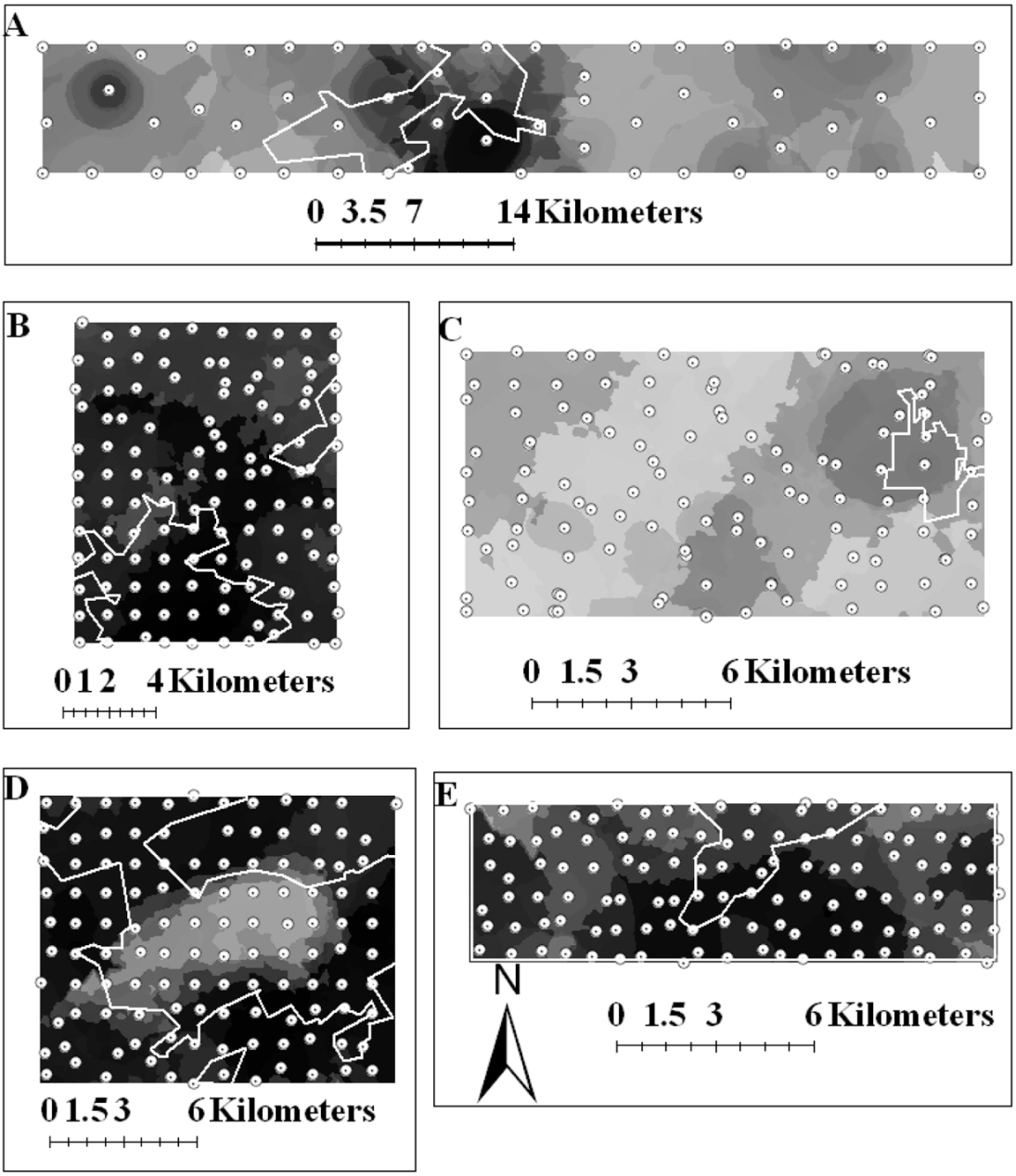

Figure 1.

Each metal's kriged estimated concentrations layered by contours (darker areas indicate higher metal concentrations based on quartile ranges) for Strips A) 1, B) 2, C) 3, D) 4, and E) 5. Layers were made semi-transparent and overlaid so dark areas indicate spatial overlap of high metal concentrations. Dots indicate sampling locations and white lines indicate approximate city locations. Strips 2, 4, and 5 had higher metal concentrations and greater overlap of metals than Strips 1 and 3.

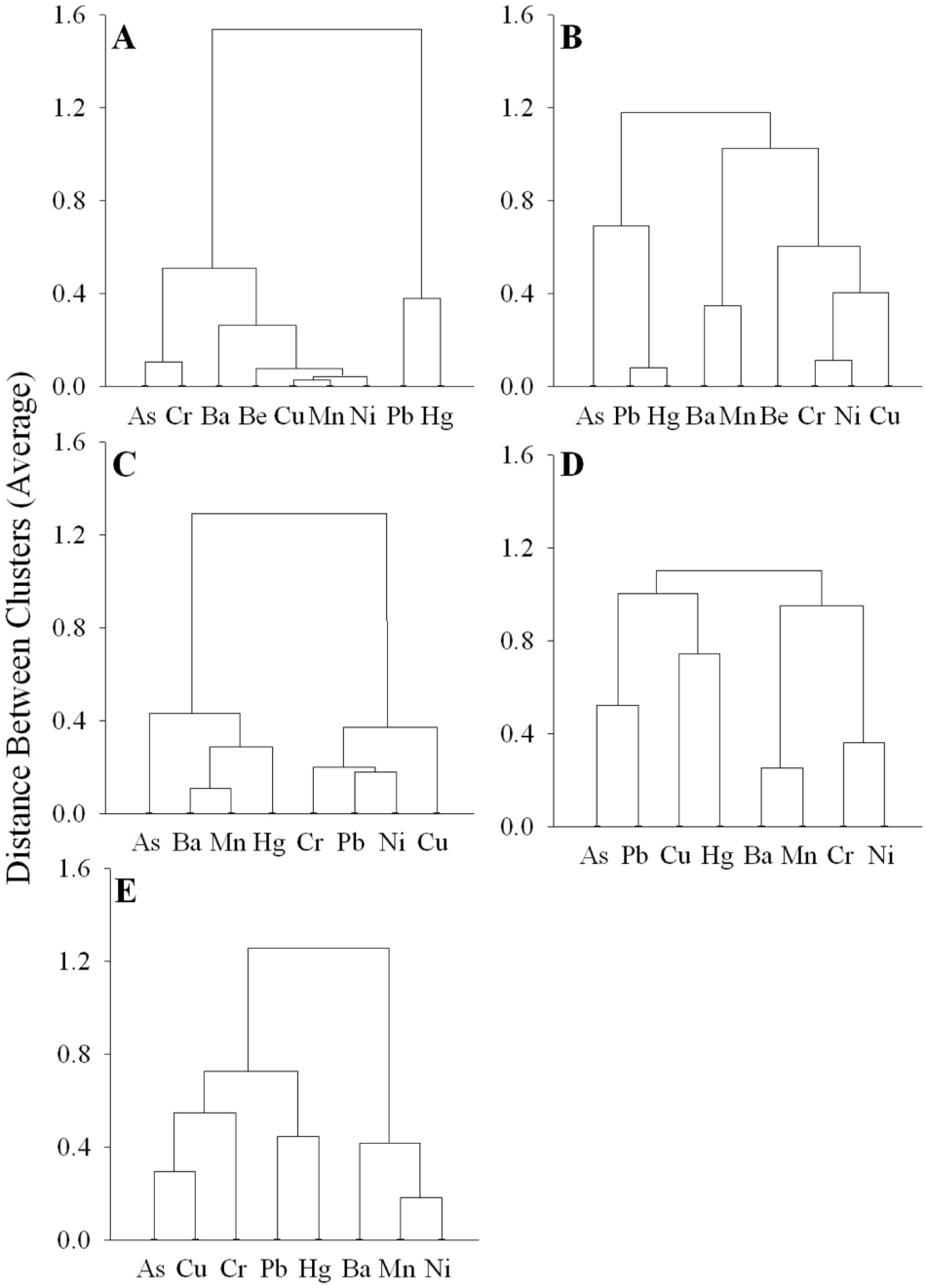

Figure 2.

Tree diagrams showing normalized root-mean-square distance between clusters for Strips A) 1, B) 2, C) 3, D) 4, and E) 5. Cluster analysis of retained factors corroborated groupings based on principal component analysis. In all strips, Ba and Mn were grouped and in Strips 1, 2, 4, and 5, Hg and Pb were grouped. Cr, Cu, Ni and As had different groupings in different strips.

Table 1

Site characteristics for Strips 1–5 compiled from health data and Federal and State GIS-based databases. Strips 2, 4, and 5 are the most similar in the majority of categories.

GIS-Based Data	Strip 1	Strip 2	Strip 3	Strip 4	Strip 5
MR/DD Rate	0.14	0.2	0.49	0.25	0.21
Population (MSA)[1]	15,000	22,000	1500	40,000 (300,000)	56,000 (300,000)
Majority Soil Types[2]	GeC, AeB, NaE, NoB, Ch, JoA	CoB, ClB2, ClC2	Ls, Cn, Wk	ClB2, ClC2	CuC, CuE
Climate Division	North central	West central	Northeast	Northwest	Northwest
Ecological Region	Sand hills	S. outer Piedmont	Carolina flatwoods	S. outer Piedmont	S. outer Piedmont
Census Land Classification	Rural	Rural	Rural	Urban	Urban
Land Cover	Forests	Forests	Forests	Urban, Built Up Land	Urban, Built Up Land
Land Use	Mixed forests	Mixed forests	Evergreen forests	Residential	Residential
TRI	3	7	1	10	2
CERCLA	9	3	1	30	13

[1] MSA: population of metropolitan statistical area (MSA) from US Census 2000

[2] GeC: Georgeville loam; AeB: Ailey sand; NaE: Nason loam; NoB: Norfolk loamy sand; Ch: Chewlaca loam; CoB: Coronaca sandy clay loam; ClB2: Cecil sandy loam, 2–6 % slope; ClC2: Cecil sandy loam, 6–10 % slope; Lf: Leaf fine sandy loam; Cn: Chipley loamy sand dark surface; Wk: Wehadkee-Chastain association; CuC: Cecil-Urban land complex, slope 2–10 %; CuE: Cecil-Urban land complex, slope 10–25 %

Table 2

Metal concentration (mg kg^{-1} dw) means and standard deviations (SD) for Strips 1–5, and ANOVA p-values for comparisons of metal concentrations between strips. Strips 2, 4, and 5 had higher mean metal concentrations than Strips 1 and 3, and all metals were significantly different between strips.

	Strip 1 Mean[a,b] (SD)	Strip 2 Mean[a] (SD)	Strip 3 Mean[b] (SD)	Strip 4 Mean[b] (SD)	Strip 5 Mean (SD)	ANOVA p-value
As	1.3 (1.3)	2.0 (2.2)	0.97 (1.0)	4.1 (5.2)	4.5 (4.6)	<0.0001
Ba	20 (29)	65 (49)	13 (7.5)	73 (60)	81 (66)	<0.0001
Be	0.13 (0.25)	0.41 (0.27)	0.004 (0.03)	0.15 (1.6)	0 (0)	0.0003
Cr	7.0 (12)	18 (16)	4.4 (4.6)	24 (19)	29 (55)	<0.0001
Cu	3.2 (4.2)	12 (13)	3.0 (7.1)	17 (19)	15 (15)	<0.0001
Pb	12 (11)	30 (31)	17 (20)	45 (50)	69 (170)	<0.0001
Mn	86 (230)	520 (610)	22 (24)	260 (180)	240 (170)	<0.0001
Ni	1.8 (3.6)	5.5 (7.0)	0.86 (1.3)	6.4 (6.8)	5.9 (6.3)	<0.0001
Hg	0.03 (0.02)	0.04 (0.03)	0.02 (0.01)	0.02 (0.03)	0.02 (0.02)	<0.0001

[a] Means reported in Aelion et al., (2008)

[b] Means reported in Aelion et al., (2009)

Table 3

Percent of variance explained by Factors 1–4 for PCA of Strips 1–5 (varimax rotation). Strip 1, 3, and 5 had only two factors retained, Strip 2 had three and Strip 4 had four factors retained based on a minimum eigen value of 1.

Strip	Factor 1	Factor 2	Factor 3	Factor 4
Strip 1	55.6 %	21.2 %	NA[a]	NA
Strip 2	27.6 %	21.9 %	17.8 %	NA
Strip 3	28 %	26.8 %	NA	NA
Strip 4	23.7 %	23.5 %	15.2 %	14.4 %
Strip 5	30.1 %	20.3 %	NA	NA

[a]Not applicable

Table 4

Communality estimates from PCA (varimax rotation) for all metals in Strips 1–5. Of all metals, As had the lowest communality estimates for Strips 1, 2, and 4, and Cr had the lowest communality estimates for Strips 3 and 5.

Metal	Strip 1	Strip 2	Strip 3	Strip 4	Strip 5
As	0.505	0.222	0.5247	0.5311	0.3616
Ba	0.8018	0.848	0.7179	0.8394	0.7624
Be	0.9112	0.7931	NA[a]	NA	NA
Cr	0.5953	0.6256	0.3976	0.8655	0.1732
Cu	0.8367	0.607	0.6119	0.673	0.4453
Pb	0.7319	0.7208	0.5429	0.6624	0.6176
Mn	0.8651	0.849	0.6124	0.857	0.6813
Ni	0.8886	0.7268	0.5592	0.9058	0.6671
Hg	0.7761	0.665	0.4147	0.8067	0.3288

[a]Not applicable