ANALYSIS OF PATTERNS IN PCDD, PCDF, AND PCB SOIL CONCENTRATIONS FROM A COMMUNITY IN MICHIGAN, USA

Chang S-C[1], Adriaens P[1], Towey T[1], Wright D[1], Demond A[1], Gillespie B[2], Franzblau A[3], Garabrant D[3]

[1]Department of Civil and Environmental Engineering, University of Michigan College of Engineering, 1351 Beal, Ann Arbor, MI 48109; [2]Department of Biostatistics, University of Michigan School of Public Health, 109 Observatory, Ann Arbor, MI 48109; [3]Department of Environmental Health Sciences, University of Michigan School of Public Health, 109 Observatory, Ann Arbor, MI 48109

Keywords: Soil, Environmental samples, North America, Patterns, PCBs, PCDD/PCDF

Introduction

The University of Michigan Dioxin Exposure Study (UMDES) was undertaken in response to concerns among the population of Midland and Saginaw Counties that the discharge of dioxin-like compounds from the Dow Chemical Company facilities in Midland have resulted in contamination of soils in the Tittabawassee River flood plain and areas of the City of Midland. There is concern that people's body burdens of polychlorinated dibenzodioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), and polychlorinated biphenyls (PCBs) may be elevated because of environmental contamination. A central goal of the UMDES was to determine the factors that explain variation in serum congener levels of PCDDs, PCDFs, and PCBs, and to quantify how much variation each factor explains. Overall study results are presented elsewhere.[1]

One potential exposure source of particular interest to the population of Midland and Saginaw counties is residential soil. In order to better understand the distribution of PCDD/Fs and PCBs in the soil of UMDES participants, an analysis of congener patterns in soil samples was performed using multivariate chemometric methods. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) were performed on the complete UMDES soil data set. The PCA and HCA output were used to generate a concentration heatmap and cluster-centroid pattern profiles. Additionally, the geographic distribution of the clusters was evaluated using a geographic information system. From the analysis, clusters emerged with distinct characteristics in congener patterns, locations, and concentration ranges.

Materials and Methods

Respondent Selection: Five populations in Midland, Saginaw, Bay, Jackson, and Calhoun Counties, Michigan, USA were sampled using a two-stage area probability household sample design. In order to be eligible for participation in the soil and vegetation sampling portion of the UMDES, subjects had to have lived in their residence at least five years and had to be the owner of their residence and property. A more detailed description of the populations and respondent selection methodology is reported elsewhere.[2]

Soil Sampling and Compositing: Each selected property was sampled in multiple locations from the surface to a depth of 6 inches. Selection of locations for sampling followed a protocol that identified the house perimeter, property areas where skin contact was likely (gardens), and areas in the flood plain of the Tittabawassee River. Each sampling station was defined by laying out a 3-foot diameter sampling ring. Three cores within the ring were collected using single-use polycarbonate tubes. The exact sampling location was recorded using a handheld global positioning system (GPS). Back in the laboratory at the University of Michigan, soils were extruded from the sealed polycarbonate tubes and the soil cores were separated into two strata: the top 1 inch (2.5cm) and bottom 5 inches

(12.5 cm). The two strata of each soil set (residence zone, soil contact, or flood plain) were combined and homogenized. Detailed soil sampling procedures can be found elsewhere.[3]

Analysis of Soil Samples: All soil samples were analyzed for the WHO designated 29 PCDD, PCDF and PCB[4] Congeners performed by Alta Analytical Laboratory, Inc. (El Dorado Hills, California, USA) by following the US EPA Method 1688A[5]and US EPA Method 8290.[6] Congeners were extracted from soil samples and quantified using high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS).

Data Cleaning: All data received from the analytical laboratory went through a data cleaning step to ensure data quality. All values below limit of detection (LOD) were replaced with the LOD divided by the square root of two ($LOD/\sqrt{2}$) to create the untransformed soil data set.

Data Transformation: Since congener data exhibited log-normal distributions, a natural logarithm transformation of ln(x+1) was undertaken. A constant-row-sum transformation was used to convert the sum of each row to unity and the natural-logarithm-transformed concentration value of each congener in each sample was converted to a fraction of unity. Finally, a range transformation was applied to each column of the dataset to ensure the variation within each congener would be similar. This final step kept the PCA from being driven by several congeners with extreme variation.[7]

Principal Component Analysis: PCA was performed using Minitab[8] software. A Scree plot, a cumulative variance plot, and principal component loading graphs were generated. The principal components that accounted for 95% of the cumulative variance were selected for further use in the HCA.[9,10]

Hierarchical Cluster Analysis: Using the selected principal components, HCA was performed based on a correlation matrix and average linkage of Euclidean distance between samples.[11] Each soil sample in the dataset was assigned a cluster membership. A dendrogram, indicating the similarity between samples and their associated clusters, was generated. Additionally, two- and three-dimensional principal component score plots - grouped by cluster membership, were produced.

Heatmap Representation: The constant-row-sum transformed data (not range transformed) were sorted according to cluster membership. Using the sorted data, a heatmap was generated to represent the congener patterns of all the soil samples in a single graph. In addition, supplemental information items, such as untransformed soil samples exceeding a particular threshold concentration, were indicated in columns adjacent to the congener pattern for each sample. From this visual representation, clear differences in the congener patterns of the clusters emerged.

Centroid Pattern Representation: Minitab software allows for the creation of a distance to cluster-centroid matrix. The sample with the smallest value for the distance to each centroid was selected to represent that cluster. The congener pattern of each cluster centroid was produced using the untransformed soil data to create 100% stacked bar graphs. Both original concentration and TEF-weighted patterns were produced.

Geographical Representation: Clustered data were projected onto a map to evaluate the geographic distribution of the clusters using ArcGIS 9.1[12] . Inferences regarding sources of congener pattern variability were made.

Results and Discussion

Results and discussion will not be available until after complete study results have been presented to the affected communities in August of 2006.

References

1. Franzblau, A, Garabrant, D, Adriaens, P, Gillespie, B, Lepkowski, J, Olson, K, Lohr-Ward, B, Ladronka, K, Sinibaldi, J, Chang, S-C, Chen, Q, Demond, A, Gwinn, D, Hedgeman, E, Hong, B, Knutson, K, Lee, S-Y, Sima, C, Towey, R, Wright, D, Zwica, L. *Organohalogen Comp* 2006 (forthcoming).
2. Olson, K, Garabrant, D, Franzblau, A, Adriaens, P, Gillespie, B, Lepkowski, J, Lohr-Ward, B, Ladronka, K, Sinibaldi, J, Chang, S-C, Chen, Q, Demond, A, Gwinn, D, Hedgeman, E, Hong, B, Knutson, K, Lee, S-Y, Sima, C, Towey, R, Wright, D, Zwica, L. *Organohalogen Comp* 2006 (forthcoming).
3. Demond A, Towey T, Chang SC, Adriaens P, Luksemburg W, Maier M, Favaro K, Wenning R, Kennington B. *Organohalogen Comp* 2006 (forthcoming).
4. Van den Berg M, Birnbaum L, Bosveld AT, Brunstrom B, Cook P, Feeley M, Giesy JP, Hanberg A, Hasegawa R, Kennedy SW, Kubiak T, Larsen JC, van Leeuwen FX, Liem AK, Nolt C, Peterson RE, Poellinger L, Safe S, Schrenk D, Tillitt D, Tysklind M, Younes M, Waern F, and Zacharewski T. *Environmental Health Perspectives* 1998;106:775-92.
5. United States Environmental Protection Agency (US EPA). Method 1668, Revision A: Chlorinated biphenyl congeners in water, soil, sediment, and tissue by high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS). Washington, DC: Office of Water, 1999.
6. United States Environmental Protection Agency (US EPA). Method 8290: Polychlorinateddibenzodioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) by high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS). Washington, DC: Office of Solid Waste and Emergency Response, 1994.
7. Johnson GW, Ehrlich R. Environmental Forensics. 2002; 3, 59.
8. Minitab, Inc., State College, Pennsylvania, USA.
9. Jolliffe IT. *Principal Component Analysis*. Springer-Verlag, New York. 1986;
10. Hair Jr JF, Black WC. Cluster analysis. In: *Multivariate Data Analysis*, Hair Jr JF, Anderson RE, Tatham RL, Black WC. (ed.), Prentice-Hall Inc, New Jersey, 1998: 147.
11.Romesburg HC. *Cluster Analysis for Researchers*. Lu-Lu Press, Carolina. 1984: 9.
12. ESRI, Inc., Redlands, CA, USA.