NIH Public Access
Author Manuscript
Environ Sci Technol. Author manuscript; available in PMC 2008 October 24.

Published in final edited form as:
Environ Sci Technol. 2008 May 15; 42(10): 3655–3661.

Geostatistical Modeling of the Spatial Distribution of Soil Dioxin in the Vicinity of an Incinerator. 2. Verification and Calibration Study

Pierre Goovaerts[1,*], **Hoa T. Trinh**[2], **Avery H. Demond**[2], **Timothy Towey**[2,#], **Shu-Chi Chang**[2,†], **Danielle Gwinn**[3], **Biling Hong**[3], **Alfred Franzblau**[4], **David Garabrant**[4], **Brenda W. Gillespie**[3], **James Lepkowski**[5], and **Peter Adriaens**[2]

1 *BioMedware Inc, Ann Arbor, Michigan 48103*

2 *Department of Civil and Environmental Engineering, University of Michigan College of Engineering, Ann Arbor, Michigan 48109*

3 *Center for Statistical Consulting and Research, University of Michigan, Ann Arbor, Michigan 48109*

4 *School for Public Health, Ann Arbor, Michigan, 48109*

5 *Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 48109*

Abstract

A key component in any investigation of cause-effect relationships between point source pollution, such as an incinerator, and human health is the availability of measurements and/or accurate models of exposure at the same scale or geography as the health data. Geostatistics allows one to simulate the spatial distribution of pollutant concentrations over various spatial supports while incorporating both field data and predictions of deterministic dispersion models. This methodology was used in a companion paper to identify the census blocks that have a high probability of exceeding a given level of dioxin TEQ (Toxic Equivalents) around an incinerator in Midland, Michigan. This geostatistical model, along with population data, provided guidance for the collection of 51 new soil data, which permits the verification of the geostatistical predictions, and calibration of the model. Each new soil measurement was compared to the set of 100 TEQ values simulated at the closest grid node. The correlation between the measured concentration and the averaged simulated value is moderate (0.44), and the actual concentrations are clearly overestimated in the vicinity of the plant property line. Nevertheless, probability intervals computed from simulated TEQ values provide an accurate model of uncertainty: the proportion of observations that fall within these intervals exceeds what is expected from the model. Simulation-based probability intervals are also narrower than the intervals derived from the global histogram of the data, which demonstrates the greater precision of the geostatistical approach. Lognormal ordinary kriging provided fairly similar estimation results for the small and

*Corresponding author phone: 734-913-1098; fax: 734-913-2201; e-mail: goovaerts@biomedware.com. Corresponding author address: BioMedware Inc, 516 North State Street Ann Arbor, MI 48104 (USA).
#Currently at LimnoTech, 501 Avis Drive, Ann Arbor, MI 48106
†Currently at National Chung Hsing University Department of Civil and Environmental Engineering, 250 Kuo Kuang Road, Taichung 40227, Taiwan

Brief
This paper uses recently collected field data to assess the accuracy and precision, then update a geostatistical model of the spatial distribution of soil dioxin around an incinerator.

Supporting Information Available Maps of summary statistics for the distribution of 100 TEQ values simulated at the level of census blocks, scatterplots of soil TEQ normal scores versus 5-year dry and wet deposition values predicted by the dispersion model, omnidirectional semivariograms of soil TEQ normal scores before and after subtracting the trend modeled from deposition data. This material is available free of charge via the internet at http://pubs.acs.org.

well-sampled area used in this validation study, however, the model of uncertainty was not always accurate. The regression analysis and geostatistical simulation were then conducted using the combined set of 53 original and 51 new soil samples, leading to an updated model for the spatial distribution of TEQ in Midland, MI.

Introduction

Air dispersion models such as ISCST3 (1) are routinely used to predict both air and soil concentrations of polychlorinated dibenzo-p-dioxins (PCDD) and dibenzofurans (PCDFs) resulting from incinerator emissions(2–4). In a companion paper (5), the ISCST3 model served as secondary information for the geostatistical prediction of soil concentrations of toxic equivalents (TEQ) emitted from an incinerator in Midland, Michigan. In particular, wet and dry depositions were used to model the spatial trend in TEQ values, leading to the mapping of the probability of exceedence of target threshold values at the census block level. The latter application allowed for a probabilistic approach to deposition mapping. A major challenge resides in the validation of these predictions, and the update of the spatial models as new information is collected.

A limited number of validation or verification studies of the ISCST3 model are available, yet they are constrained by (i) the small number of samples available, or (ii) sampling locations that are biased towards the local, downwind locations expected to exhibit high concentration values. For example, in a study of a municipal waste incinerator in Columbus, Ohio, Lorber et al. (2) validated their prediction model using five ambient air sites and 34 soil samples taken at a distance of about 8 km away from the incinerator over the two reference years of 1995–1996. They observed that both soil- and air-predicted concentrations were within a factor of 10 of observations. The major uncertainties affecting predictions were: source characterization (stack emission rates of dioxins); meteorological data; and atmospheric and soil fate and transformation processes of the dioxins. In another study conducted in France (6), 75 soil samples were collected during an extensive sampling campaign. In simple terrain they found a significant association between modeled dioxin ground-level air concentrations and measured dioxin soil concentrations. However, in a complex topography situation, the model overpredicted ground-level air concentrations. Other validation studies often have less than ten samples each for air and soil media (4,7,8), and predictions are again in the order of magnitude range.

It is not common for models to be updated or further modified to incorporate the results from validation studies. Several studies have modified the ISCST3 model to account for partitioning of the contaminant between vapor and particle phase (9) or included a simple soil mixing model considering the decay of dioxins in the soil reservoir (2). The former modification led to more accurate predictions of soil concentration for simple and flat terrains, but resulted in overprediction for complex terrains. When including a soil mixing model, predicted values were lower than measured concentrations for the tetra and penta CDD/Fs while the opposite trend was observed for the hexa to octa CDD/Fs. Predicted concentrations were within a factor of 2 for total soil TEQ.

Although the ISCST3 model has been used and modified considerably, the validation process is still rather crude. For studies with few observations, the validation takes the form of a table with columns of predicted and measured values or a plot of the two sets of prediction and observation data (4,7,8). When more observations are available, contour maps of predicted and measured soil concentrations were compared (2). Not much attention has been paid to evaluate and incorporate the local uncertainties of the model predictions in the validation study. Since the ISCST3 requires substantial data inputs not readily available at certain locations (e.g. longterm average meteorological data, wind speed and direction, stack emission, landscape characterization, and particle size distribution), probabilistic approaches were used to propagate the uncertainty associated with selected input data (5). To address the need to incorporate spatial complexity of soil data, geostatistical modeling that incorporates the spatial coordinates of field data is a useful tool. The underlying probabilistic approach of the model would take into account the model uncertainties (10,11). This approach allows for many advanced validation techniques to quantify the accuracy and precision of model predictions (12).

In this paper, we present (i) a validation approach for the ISCST3- informed predictions of spatially-distributed TEQ in soils around the Midland, Michigan incinerator using data collected under the University of Michigan Dioxin Exposure Study, and (ii) the results from an updated geostatistical model that integrates newly collected data and updated calibration of ISCST3 predictions and soil concentrations in the Midland area.

Materials and Methods

Study Area

The study area is the vicinity of the Dow Chemical Company facility in Midland, Michigan (Figure S1, Supporting Information). Accounting for 53 field data and the output of the EPA Industrial Source Complex (ISC) dispersion model, the spatial distribution of dioxin around the Dow facility incinerator was modeled using geostatistical simulation in the companion paper (5). The set of simulated maps were used to compute, for each census block, the probability that the average TEQ value within that block exceeds a threshold of 90 ppt, which is the generic soil residential Direct Contact Criterion (DCC) used by the Department of Environmental Quality for Midland, Michigan (13). Census blocks that were the most likely to exceed the DCC threshold and have the largest population at risk were targeted for a recent soil sampling campaign. This estimation was based on using a conservative threshold of 75 ppt TEQ (Figure S1c, Supporting Information), which yielded 196 census blocks with a probability greater or equal to 0.4. This campaign was conducted within the framework of the University of Michigan Dioxin Exposure Study (UMDES) that focuses on quantifying exposure pathways to dioxins from industrial sources relative to background exposures.

Soil Sampling and Measurement

The UMDES soil sampling campaign was carried out at 51 locations in the vicinity of the Dow Chemical's incinerator complex. Due to confidentiality agreements, the sampling locations can not be disclosed. The soil sampling and analysis are described in Demond et al. (14). Analytical results of the 1-inch depth UMDES soil samples provide concentrations of 29 dioxins, furans and PCB congeners. Total PCDD/PCDF, excluding PCBs' contributions, of these 51 UMDES samples were used to validate the deposition model.

Validation of the Geostatistical Model of Uncertainty

The geostatistical model consists of 100 maps of TEQ values simulated on a 261×261 grid with a 50 m grid spacing. The set of simulated values is denoted $\{z^{(l)}(\boldsymbol{u}_j);\ l=1,\ldots, L\ ;\ j=1,\ldots, N\}$, with L=100 and N=$(261)^2$. Each of the 51 new observations, $\{z(\boldsymbol{u}_\alpha),\ \alpha=1,\ldots, n\}$, was compared to the distribution of 100 TEQ values simulated at the closest grid node $\boldsymbol{u}_i$, with $\|\boldsymbol{u}_i - \boldsymbol{u}_\alpha\| < \|\boldsymbol{u}_j - \boldsymbol{u}_\alpha\|\ \forall\ j$. The coordinates of these 51 samples were unavailable at the time of the creation of the air dispersion model and the generation of simulated TEQ maps in the companion paper. In addition, the average distance between the new observations and the closest grid node is 17.6 m, which is negligible relative to the range of autocorrelation of 776 m. The validation stage proceeded as follows:

NIH-PA Author Manuscript

NIH-PA Author Manuscript

1. Boxplots were used to visualize where the measured value $z(\boldsymbol{u}_\alpha)$ falls within the distribution of 100 simulated values, $\{z^{(l)}(\boldsymbol{u}_i);\ l=1,\dots,100\}$.

2. From the distribution of 100 simulated values we computed, for each of the 51 new sampled locations, a series of symmetric *p*-probability intervals (PI) bounded by the (1-*p*)/2 and (1+*p*)/2 quantiles of that distribution. For example, the 0.5-PI is bounded by the lower and upper quartiles. A correct modeling of local uncertainty would entail that there is a 0.5 probability that the actual TEQ value at that location falls into that interval or, equivalently, that over the study area 50% of the 0.5-PI include the true value. The fraction of true values falling into the symmetric *p*-PI is estimated as:

$$\bar{\zeta}(p)=\frac{1}{n}\sum_{\alpha=1}^{n}\zeta(\mathbf{u}_\alpha;p) \tag{1}$$

   where $\zeta(\mathbf{u}_\alpha;\ p)$ equals 1 if $z(\boldsymbol{u}_{\alpha})$ lies between the (1-*p*)/2 and (1+*p*)/2 quantiles, and zero otherwise. The scattergram of the estimated, $\bar{\zeta}(p)$, versus expected, *p*, fractions is called the “accuracy plot” and quantifies the accuracy of the model of uncertainty (11,12). The average width of the PIs that include the new observations informs on the precision of the models of local uncertainty.

3. The correlation was computed between each measured TEQ value $z(\boldsymbol{u}_\alpha)$ and the average simulated value $\bar{z}(\mathbf{u}_i)$ calculated as:

$$\bar{z}(\mathbf{u}_i)=\sum_{l=1}^{100}z^{(l)}(\mathbf{u}_i) \tag{2}$$

4. Prediction errors, $\bar{z}(\mathbf{u}_i)-z(\boldsymbol{u}_\alpha)$, were mapped to identify any spatial pattern for the over and under-estimation of TEQ values.

The same validation was conducted on the estimates based on lognormal kriging, substituting the 1000 backtransformed quantiles $\varphi^{-1}(y_p(\mathbf{u}_i))$ (recall eq S11 in the companion paper) for the 100 simulated TEQ values $z^{(l)}(\boldsymbol{u}_i)$.

Updating of the Geostatistical Model

The procedure used to derive the first model of uncertainty and described in detail in the companion paper (5) was applied to the combined set of 53 data collected during 1983–1998 sampling campaigns and the new 51 observations from the UMDES sampling. The methodology follows:

1. The 104 soil PCDD/PCDF TEQ concentrations were normal score transformed to correct for the strongly positively skewed sample histogram. Because the UMDES samples were preferentially located in census blocks with expected high level of dioxins, sample statistics likely overestimate the magnitude of the contamination over the entire modeled area. In order to reduce the weight assigned to these redundant observations and obtain statistics (e.g. mean, variance) that are more representative of the distribution of TEQ values within the modeled area, declustering weights were computed and used in the normal score transform; recall equation (S1, Supporting Information) in companion paper (5). A cell-declustering method, with cell sizes ranging from 50 to 350 meters, was applied and the cell size of 218 meters leading to the smallest declustered mean was retained since high values were preferentially sampled (15).

2. The 104 transformed data were regressed against the deposition (wet and dry) values predicted using the numerical dispersion model (base case scenario). This regression model was used to predict the TEQ concentration and standard error at the nodes of

the 261×261 simulation grid centered on the incinerator. The grid, which has a 50 m spacing, does not include any node within the boundary of the plant.

3. The spatial variability of regression residuals was modeled using the semivariogram.

4. Sequential Gaussian simulation was used to simulate the spatial distribution of TEQ values conditionally to the 104 soil TEQ data, the trend model inferred from the calibration of the deposition data (step 2) and the pattern of correlation modeled in step 3. One hundred realizations were generated over the 261×261 simulation grid.

5. Point simulated values were averaged within each census block to yield a simulated block value (upscaling). This averaging is repeated for each realization, yielding a set of 100 simulated values for each census block. The following three statistics were derived from the distribution of 100 simulated block TEQ values: mean, variance and proportion of block values that exceeds a threshold of 90 ppt which is the soil generic residential Direct Contact Criterion (DCC) used by the Department of Environmental Quality for Midland, Michigan (13).

The procedure SAS GLM (16) was used for the regression, while the declustering and normal score transform were conducted using the programs *declus* and *nscore* in the public domain software library, GSLIB (15). Sequential Gaussian simulation with local means was implemented by modifying the FORTRAN source code Sgsim in GSLIB. Aggregation within census blocks and mapping were accomplished using the commercial product, TerraSeer STIS (Space-Time Intelligence System) (17).

Results and Discussion

Previously-available versus UMDES TEQ data

Figure 1 (a, b) shows the histograms of the 53 TEQ data used to build the geostatistical model described in Goovaerts et al. (5) and the 51 TEQ data collected within the framework of the University of Michigan Dioxin Exposure Study (UMDES). UMDES data exhibit larger mean and median values, while their maximum (923 ppt) is more than twice the maximum of 450 ppt observed in the dataset used for the original model. This observation underscores the spatial heterogeneity of dioxin deposition, considering that the first dataset (EPA, MDEQ (13)) was collected in areas near the plant with expected high values, while the 51 new UMDES data were preferentially located in census blocks with TEQ values in excess of 75 ppt. The geographic distribution of the sampled census blocks includes locations North, East and South-East of the incinerator complex (Figure S1a, b). The data statistics may further indicate the contribution of multiple sources of dioxins in soil, aside from incinerator deposition; preliminary fingerprinting analysis of the deposition patterns indeed indicates a divergence of some samples based on their congener contributions to the TEQ (data not shown; (18)).

In addition to the differences in summary statistics, the two normal score transformed datasets correlate differently with the five-year dry and wet deposition values predicted by the dispersion model. While the 53 old data were better correlated with dry deposition ($r=0.641$) than wet deposition ($r=0.344$) (5), the opposite pattern is displayed by the 51 UMDES data. The strongest correlation ($r=0.691$) is observed for wet deposition, while dry deposition has a smaller correlation ($r=0.513$) with the data (Figure S2). The correlation coefficients are significantly different for wet deposition ($p=0.015$) but non-significant for dry deposition ($p=0.339$). This change in the pattern of correlation can be explained by the fact that the South-western side of the plant, where most of wet deposition occurred according to the numerical model, was not sampled during the UMDES campaign because of its low population density (and thus decreasing likelihood to be included in a random census block based sampling design). Previous studies, which did not employ the ISCST3 model, indicated the importance of a regional component in designating whether wet or dry deposition were dominant for dioxin deposition (e.g. (19–21)). Dry deposition tended to dominate in the Midwest; wet deposition was most important in Western Europe and the Houston area.

Validation study

Each of the 51 UMDES data was compared to the distribution of 100 TEQ values simulated at the closest grid node. Box plots in Figure 2 summarize the results for all 51 sampled locations. The UMDES-sampled TEQ values tend to fall within the upper tail of the simulated local distributions (based on EPA/DEQ data, (5): 42 observations out of 51 are above the median, with 7 data exceeding the 95 percentile. This underestimation by the geostatistical model is confirmed by the scatter plot of observed concentrations versus averaged simulated values (Figure 3a). Although predicted values are on average lower than observed values (44.24 versus 62.29 ppt), there is a reasonably good agreement between the two datasets (correlation=0.44). Note that these statistics are computed after discarding the extreme observation of 923 ppt. The difference in summary statistics between the dataset used for the prediction model and those collected for UMDES explains this underestimation of concentrations. The map of prediction errors (not shown for confidentiality reasons) indicates that the underestimation occurs mainly in the vicinity of the plant property line.

Lognormal kriging displays a similar underestimation (Figure 3b), although the use of 8 neighbors instead of 32 for the simulation procedure (Supporting Information in companion paper) reduces the smoothing effect. Figure 3c indicates that both average simulated and estimated values are fairly similar. Indeed, discrepancies between the two approaches, recall Figure S2 in (5), are observed farther away from the incinerator, where UMDES data were not collected. Differences between approaches are larger in terms of uncertainty assessment: the average kriging standard error (51.1) exceeds the average standard deviation of the local distribution of 100 simulated values (44.4). Yet, the spatially confined area analyzed in the UMDES study does not allow for a realistic assessment of the benefit of incorporating the air dispersion model which is expected to increase away from the incinerator (e.g. SE corner) where no field data are available (Figure S2b, Supporting Information in (5)).

One consequence of the underestimation of TEQ concentrations is that only one averaged simulated value or two kriged values exceed the residential Direct Contact Criterion (DCC) of 90 ppt, see Figure 3a (upper left quadrant). Furthermore, only one of the 13 UMDES data exceeding 90 ppt is actually reflected in the geostatistical model (Figure 3b, upper right quadrant). This is due to the fact that all least-square interpolation algorithms tend to underestimate the high values while the low values are overestimated (e.g. (22). Considering that this study attempts to provide a probabilistic interpretation of spatially differentiated data, the interpretation needs to be supplemented by a measure of the likelihood that the soil generic residential Direct Contact Criterion (DCC) of 90 ppt is exceeded. This likelihood was computed, for each simulation grid node, as the proportion of 100 simulated TEQ values that are larger than 90 ppt (Figure 4e in (5)). The average likelihood is 0.12 for the 38 UMDES data below 90 ppt and 0.21 for the 13 UMDES data above 90 ppt. The analysis of lognormal kriging results yields similar probabilities of 0.1 and 0.2, respectively. Despite the low average probability of exceedence resulting from the underestimation of TEQ concentrations, the geostatistical measure of uncertainty allows discrimination between locations with potentially high and low dioxin levels. This result also illustrates the potential hazard of ignoring the uncertainty attached to estimates in the decision-making process (23, 24).

Figure 3e shows that the simulation-based model of uncertainty is accurate: the proportion of observations that fall within probability intervals (PI) exceeds what is expected from the model. For example, the observed TEQ value is included in the 0.5-PI for 59% of the 51 new samples (expected proportion=50%). The accuracy of the kriging-based model fluctuates with the probability p; the PIs contain a smaller than expected fractions of true values for probabilities larger than 0.45. Not only should the true TEQ value fall into the PI according to the expected probability p, but this interval should be as narrow as possible to reduce the uncertainty about that value. The average width of these local PIs should also be smaller than the global PI inferred from the sample histogram. The scatter plot in Figure 3f indicates that, for both approaches and all probabilities p, the local PIs are narrower than the corresponding global PIs, which means that the geostatistical model of uncertainty is more precise than an aspatial model that ignores the location of soil samples.

Updated geostatistical model

The creation of an updated geostatistical model started with the regression of the output of the air dispersion model ISCST3 (dry and wet depositions plus their interaction) against the set of 104 normal score transformed TEQ data (to include both the previously collected EPA/DEQ data and the UMDES data). As for the 53 old data, TEQ values are better correlated with dry deposition (r=0.511) than wet deposition (r=0.361), although the difference is not significant (p=0.181). Application of the regression model to the 261×261 exposure assessment grid yields the trend model displayed in Figure 4a. The joint contribution of dry and wet depositions generates a continuous ring of high trend values around the plant. Comparison of the old and updated trend models in Figure 4b shows a decline in regression estimates on the southwestern side of the plant where wet deposition is the predominant mechanism, while the regression leads to larger estimates downwind of the plant where dry deposition prevails. These differences reflect changes in the calibration of the trend model which likely result from the more geographically distributed census-block based sampling of UMDES data: none of the 51 additional samples was collected South of the plant because of the low population density.

The incorporation of 51 new data changed the shape of the experimental semivariogram of normal score TEQ values that is now clearly bounded, with a range of autocorrelation of around 4 km (Figure S3a). The trend component and semivariogram model were used in sequential Gaussian simulation to generate 100 realizations of the spatial distribution of TEQ values. Figure 4c–e shows the maps of the mean and standard deviation of the distribution of 100 TEQ values simulated at the level of census blocks. The difference map of Figure 4d indicates that, except for a slight decline on the South-Eastern side of the plant, the new model displays larger TEQ values close to the plant, in particular in the North and South. This trend is consistent with the results of the validation procedure that revealed an underestimation by the old geostatistical model that occurred mainly in the vicinity of the plant property line. While the larger predicted values North of the plant reflect the influence of UMDES samples, smaller estimates on the Southern side are caused by the new calibration of the air dispersion model (recall Figure 4b).

The proportional effect is a particular form of heteroscedasticity where the local variance of data is related to their local mean (24,25). This effect explains the similarity between the spatial pattern of the maps of mean and standard deviation (Figure 4c–e): the rank correlation coefficient between the two sets of block-level values is 0.97. A decline in standard deviation, which translates into smaller uncertainty about block-level TEQ values, is also observed in the Northern and Eastern parts of the study area where most of the UMDES samples were collected (Figure 4f). Further calibration of the model requires incorporation of topographical changes (e.g. 5), temporal information on the climate and particle density variability that impacts local vs. distal deposition (26), and exclusion of samples whose homolog or congener profile does not support their origin as incinerator point sources (e.g. (18)).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Acknowledgements

Financial support for this study comes from the Dow Chemical Company through an unrestricted grant to the University of Michigan. The authors acknowledge Ms. Sharyn Vantine for her continued assistance and Drs. Linda Birnbaum, Ron Hites, Paolo Boffetta and Marie Haring Sweeney for their guidance as members of our Scientific Advisory Board. The first author's contribution was partially funded by grant R44-CA105819-02 from the National Cancer Institute. The views stated in this publication are those of the author and do not necessarily represent the official views of the NCI.

Literature Cited

1. U.S. Environmental Protection Agency. User's Guide to the Industrial Source Complex Dispersion Models Vol. 2 Description of Model Algorithms, EPA-454/B-95-003b. USEPA; 1995.
2. Lorber M, Eschenroeder A, Robinson R. Testing the USA EPA's ISCST-Version 3 model on dioxins: a comparison of predicted and observed air and soil concentrations. Atmospheric Environment 2000;34 (23):3995–4010.
3. Lohman K, Seigneur C. Atmospheric fate and transport of dioxins: local impacts. Chemosphere 2001;45(2):161–171. [PubMed: 11572608]
4. Oh JE, Choi SD, Lee SJ, Chang YS. Influence of a municipal solid waste incinerator on ambient air and soil PCDD/Fs levels. Chemosphere 2006;64(4):579–587. [PubMed: 16410018]
5. Goovaerts P, Adriaens P. Geostatistical modeling of the spatial distribution of soil dioxin in the vicinity of an incinerator. 1. Theory and application to Midland, Michigan. Environ Sci Technol. 2007Submitted
6. Floret N, Viel JF, Lucot E, Dudermel PM, Cahn JY, Badot PM, Mauny F. Dispersion modeling as a dioxin exposure indicator in the vicinity of a municipal solid waste incinerator: A validation study. Environmental Science & Technology 2006;40(7):2149–2155. [PubMed: 16646446]
7. Chao MR, Hu CW, Chen YL, Chang-Chien GP, Lee WJ, Chang LW, Lee WS, Wu KY. Approaching gas-particle partitioning equilibrium of atmospheric PCDD/Fs with increasing distance from an incinerator: measurements and observations on modeling. Atmospheric Environment 2004;38(10): 1501–1510.
8. Yoshida K, Ikeda S, Nakanishi J, Tsuzuki N. Validation of modeling approach to evaluate congener-specific concentrations of polychlorinated dibenzo-p-dioxins and dibenzofurans in air and soil near a solid waste incinerator. Chemosphere 2001;45(8):1209–1217. [PubMed: 11695634]
9. Basham JP, Whitwell I. Dispersion modelling of dioxin releases from the waste incinerator at Avonmouth, Bristol, UK. Atmospheric Environment 1999;33(20):3405–3416.
10. Saito H, Goovaerts P. Accounting for measurement error in uncertainty modeling and decision-making using indicator kriging and p-field simulation: application to a dioxin contaminated site. Environmetrics 2002;13(5–6):555–567.
11. Saito H, Goovaerts P. Selective remediation of contaminated sites using a two-level multiphase strategy and geostatistics. Environmental Science & Technology 2003;37(9):1912–1918. [PubMed: 12775065]
12. Deutsch, CV. Direct assessment of local accuracy and precision. In: Baafi, EY.; Schofield, NA., editors. Geostatistics Wollongong '96. Kluwer Academic Publishers; Dordrecht: 1996. p. 115-125.
13. Michigan Department of Environmental Quality. [Accessed July 2]. Available: http://www.deq.state.mi.us/documents/deq-whm-hw-dow-2005-06-10-midland-summary-map.pdf
14. Demond A, Adriaens P, Towey T, Chang S-CBH, Chen Q, Chang C-W, Franzblau ADG, Gillespie BEH, Knutson K, Lee C, Lepkowski J, Olson K, Ward B, Zwica L, Luksemburg W, Maier M. Residential Soil Concentrations of PCDDS/PCDFS/PCBS in a Community in Michigan. Environ Sci Technol. 2007In Review
15. Deutsch, CV.; Journel, AG. GSLIB: Geostatistical Software Library and User Guide. Oxford University Press; New York: 1998.
16. SAS Institute Inc. SAS/STAT User's Guide. 4. 2. SAS Institute Inc; Cary, NC: 1989.
17. Avruskin GA, Jacquez GM, Meliker JR, Slotnick MJ, Kaufmann AM, Nriagu JO. Visualization and exploratory analysis of epidemiologic data using a novel space time information system. Int J Health Geogr 2004;3(26)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

18. Trinh H, Adriaens P. Congener-Specific Differentiation of Soil Samples in the City of Midland, Michigan Using Multidimensional Scaling. Organohalogen Compounds. 2007
19. Eitzer BD, Hites RA. Atmospheric transport and deposition of polychlorinated dibenzo-p-dioxins and dibenzofurans. Environ Sci Technol 1989;23(11):1396–1401.
20. Koester CJ, Hites RA. Wet and Dry Deposition of Chlorinated Dioxins and Furans. Environmental Science & Technology 1992;26(7):1375–1382.
21. Schroder J, WelschPausch K, McLachlan MS. Measurement of atmospheric deposition of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) to a soil. Atmospheric Environment 1997;31(18):2983–2989.
22. Goovaerts, P. Geostatistics for Natural Resources Evaluation. Oxford University Press; New York: 1997.
23. Barabas N, Goovaerts P, Adriaens P. Geostatistical assessment and validation of uncertainty for three-dimensional dioxin data from sediments in an estuarine river. Environmental Science & Technology 2001;35(16):3294–3301. [PubMed: 11529567]
24. Saisana M, Dubois G, Chaloulakou A, Spyrellis N. Classification criteria and probability risk maps: Limitations and perspectives. Environmental Science & Technology 2004;38(5):1275–1281. [PubMed: 15046326]
25. Journel, AG.; Huijbregts, CJ. Mining Geostatistics. Academic Press; London: 1978.
26. Lohmann R, Jones KC. Dioxins and furans in air and deposition: A review of levels, behaviour and processes. Science of the Total Environment 1998;219(1):53–81. [PubMed: 9770324]

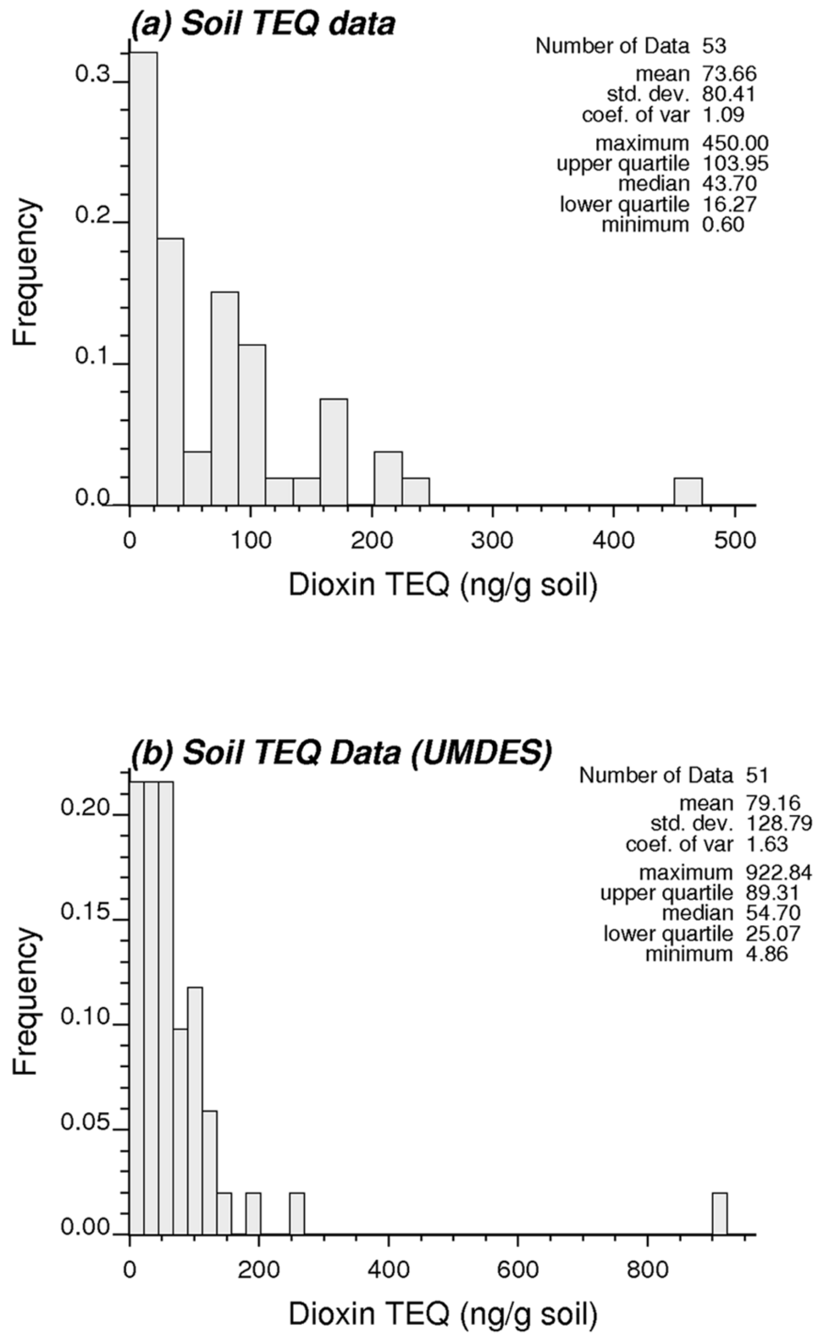

Figure 1.

Histograms of: (a) 53 TEQ data used to construct the original geostatistical model and (b) the 51 new samples collected during the UMDES campaign.

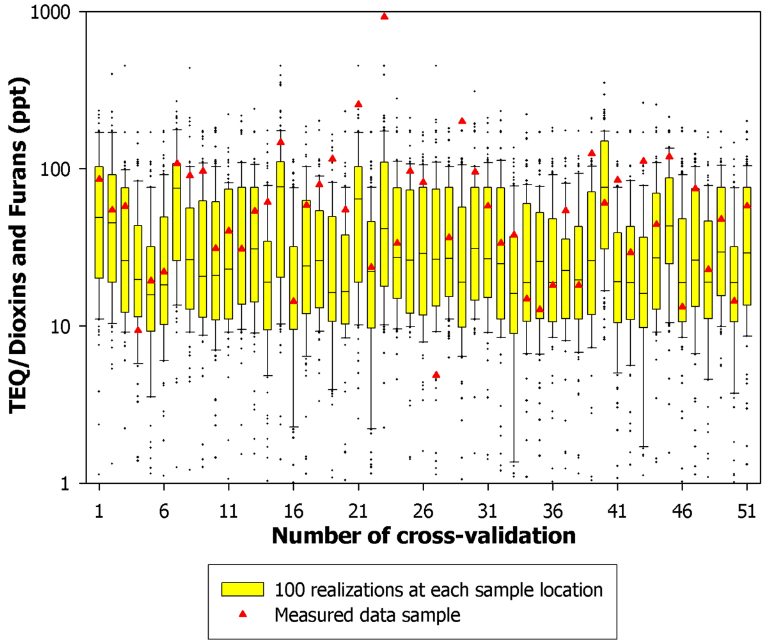

Figure 2.

Box plots of the distributions of 100 simulated TEQ values corresponding to the grid nodes the closest to the location of the 51 UMDES samples (values denoted by red triangles).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

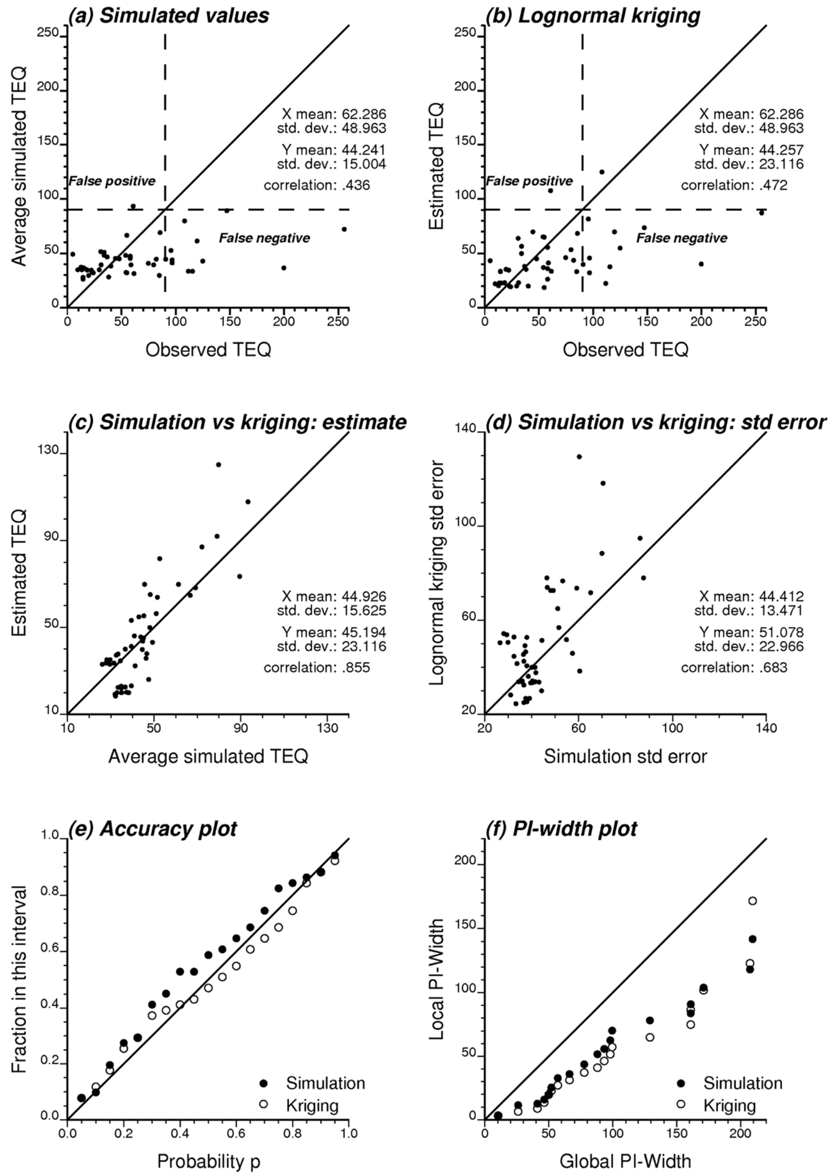

Figure 3.

Scatterplot of UMDES observations versus the mean of the 100 simulated TEQ values (a) or the lognormal kriging estimate (b) (the maximum observation of 923 ppt is not included for graph clarity). Scatterplots of simulation versus lognormal kriging results (c,d). Plot of the proportion of observed TEQ values falling into probability intervals (PI) of increasing size (accuracy plot, e). The width of these local PIs is plotted against the width of the global PIs that are derived from the sample histogram (f).

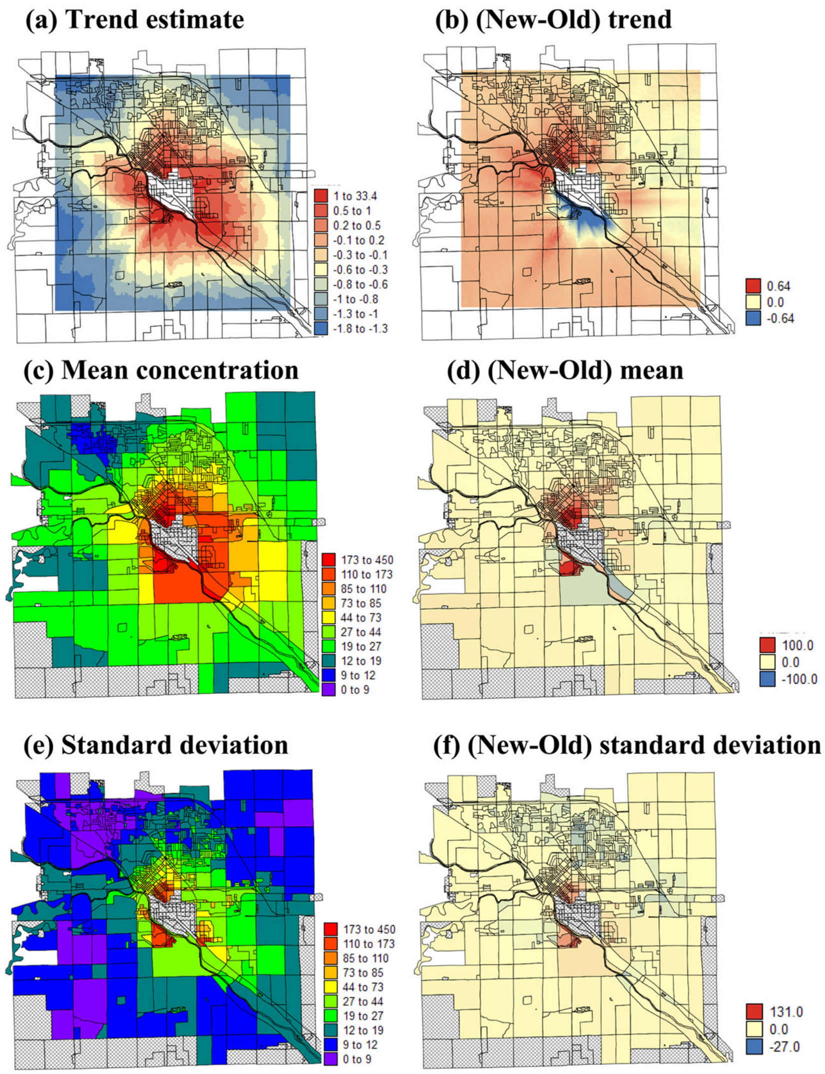

Figure 4.

Impact of UMDES data on the definition of the trend component (a,b), and the mean (c,d) and standard deviation (e,f) of the distribution of 100 TEQ values simulated at the level of census blocks. Left column shows the results for the updated geostatistical model, while the right column illustrates differences with the old model. Hatched polygons denote census blocks outside the simulation area.