M-Scale Model for Multi-Scale Estimation of Dioxin Data

Meng-Ying Li[1], Noemi Barabas[2], Peter Adriaens[1]

[1]University of Michigan
[2]Limno-Tech, Inc.

Introduction

Reliable characterization of the spatial distribution of sediment site attributes, such as dioxin concentrations or the impact of microbial activity on dioxin patterns depends on how well sampled values represent all values throughout the entire study site. Whereas geostatistical tools have been developed to interpolate the attribute values in space, these do not explicitly take into account the uncertainties associated with the various scales (field, lab, mesocosm) at which the data have been collected. Here, we describe a recently developed statistical model (hereafter M-Scale model) to optimize the reliability of sampled data on a multi-scale basis. The model not only serves as a tool to evaluate parameter relationships over different scales by their covariances and data uncertainty, but also make further use of these covariances and data uncertainty as basis for a precision-optimized estimator. Information from each scale will be weighted by the projected similarity to the scales of interest, with adjustments considering the different precision they provide. Unlike conventional geostatistic tools that are based on the point-to-point spatial structures, the multi-scale model introduces a new framework for spatial analysis in which regional values at different scales are anchored by the correlations of each other.

Materials and Methods

The Passaic River (New Jersey) is contaminated with, among others, polychlorinated dibenzo-p-dioxins (PCDD) derived from a multitude of sources. To aid in remedial decision-making, research conduced over the last decade has attempted to quantify the distribution of dioxins in the river.[1]

To address the impact of scale on uncertainty estimations about collected variables (here: PCDD measurements), the principles developed for a previously published multi-resolution model[4] were adopted to develop the M-Scale data estimator. The statistical premise of this model is that the variation of values at different spatial scales is heuristically related. In other words, if samples are collected at the “point scale” (e.g. single cores selected from large sites), the variation of parameters collected within the core are propagated to higher scales by taking into account the correlations of the local averages between scales, as shown in Figs. 1 and 2. These correlations are informed by the variation of local averages at each scale. Taken together, the contributions of variation at each scale are weighted and aggregated into a best linear unbiased estimator, as shown in Fig. 3.

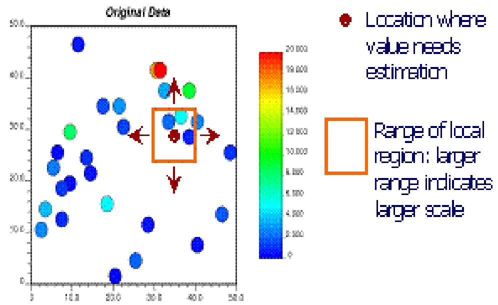

Fig.1 Calculating Local Average Values

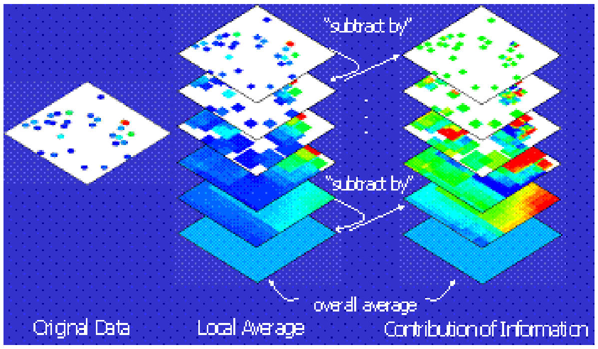

Fig.2 Evaluating Contribution of Variation from Different Scales

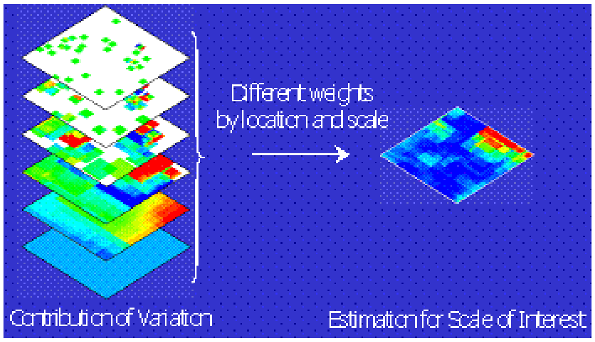

Fig.3 Best Linear Unbiased Estimator by Weighting Variations from Different Scales

The overall goal of this study is to assess the estimation performance by Ordinary Kriging and M-Scale techniques in spatial estimation to a linearized representation of the dioxin dataset.[2]

The specific objectives are to:

1. Apply the estimation techniques to three xz-planes (vertical transect planes along the river flow) to visually compare the spatial features sustained by each model.

2. By the result of cross-validation, compare the reproduction of estimation against true data on the basis of data amounts in each magnitude range.

Cross-validation in the second objective refers to the approach that each datum is sequentially drawn from the original data set and compared to the estimate of this datum by the rest of the data.

Results and Discussion

Original tetra-chloro-dibenzo-p-dioxins (TCDD) measurements collected in the Passaic River (1996 dataset) are plotted as the basis of visual comparison against the estimation maps, as shown in Fig. 4.

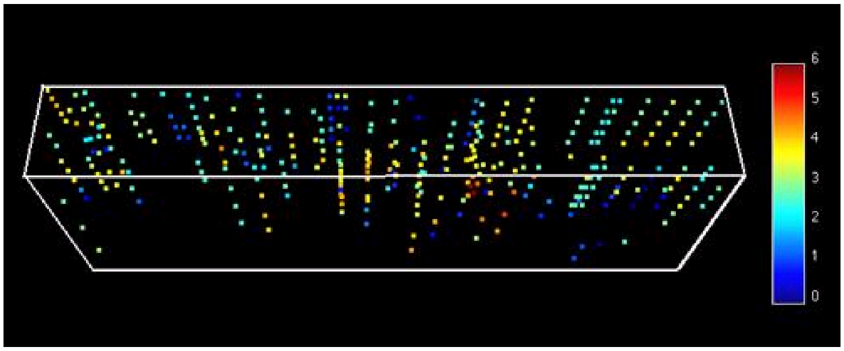

Fig. 4 Original TCDD Measurements Plotted at Their 3-D Coordinates

The preliminary results are represented as follows:

1. A comparison of Ordinary Kriging estimate (left) with M-Scale estimation (right) for three selected cross-sections of the river (Fig. 5).

2. Quantile-quantile (Q-Q) plots[3] from both models of the estimated chlorination in space (Fig. 6).

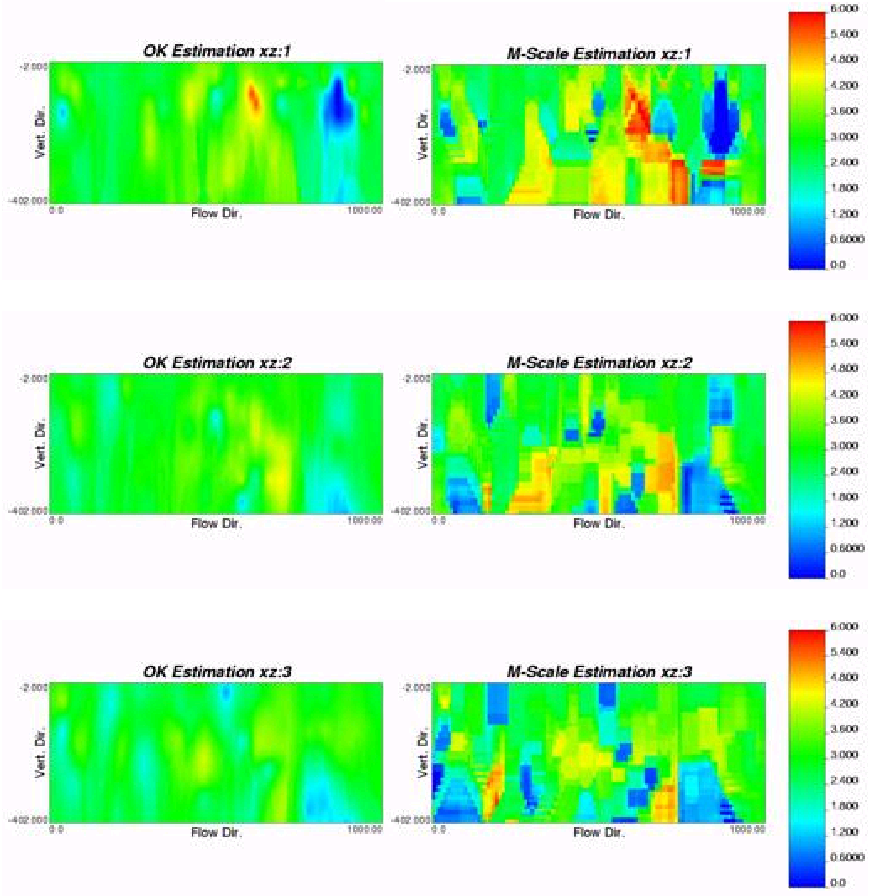

Fig. 5 Estimation Maps by Ordinary Kriging (OK) and M-Scale Models

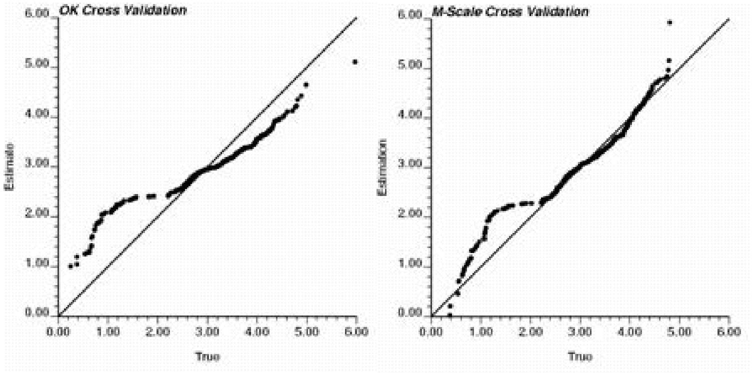

Fig. 6 Q-Q Plots from Ordinary Kriging (OK) and M-Scale Models

Estimation maps from the two models indicate the following:

1. Ordinary Kriging estimation smoothens out the “neighborhood” of important chlorination amounts.

2. The M-Scale model preserves and incorporates point measurements through the scaling process and retains important contributions and features, especially in the regions where hot-spot measurements exist.

The Q-Q plot for the cross-validation of estimates represents graphically how well model estimates predict the actual values (deviation from the 90° line). The results further confirm that the M-Scale estimation generally reflects the data across the entire magnitude range, while Ordinary Kriging tends to underestimate higher values and overestimate lower values.

Acknowledgements

The authors acknowledge the DODDOE-EPA Strategic Environmental Research and Development Program (SERDP) for funding this work.

References

1. Barabas, N., P. Goovaerts, and P. Adriaens (2001). Geostatistical Assessment and Validation of Uncertainty for Three-Dimensional Dioxin Data from Sediments in an Estuarine River, Environ. Sci. Technol. Environ. Sci. Technol. 35: 3294-3301.

2. Barabas, N., P. Goovaerts, and P. Adriaens (2004). Modified Polytopic Vector Analysis to Identify and Quantify Dioxin Dechlorination Signatures in Sediments. 2. Application to the Passaic River Superfund Site, Environ. Sci. Technol., 38: 1821-1827.

3. Goovaerts, P. (1997). Geostatistics for Natural Resources Evaluation, Oxford University Press, ISBN 0-19-511538-4.

4. Wilsky, A. S. (2002). Multiresolution Markov Models for Signal and Image Processing, Proceedings of the IEEE 90:8 1396-1458.