BOREHOLE AND MICROTREMOR DATA FUSION FOR SEDIMENT LAYER THICKNESS ESTIMATION IN BEIJING AREA

Eric Rohrbach[1], Lanbo Liu[1,2], Qifu Chen[2], and Weijun Wang[2]

[1]*Dept. of Civil & Environmental Engineering, University of Connecticut, Storrs, CT 06269, USA*

[2]*Institute of Earthquake Science, China Earthquake Administration, Beijing 100036, PRC*

Eric.Rohrbach@Uconn.edu, Lanbo.Liu@UConn.edu, wjwang@seis.ac.cn

ABSTRACT

A method for synergizing the information on sedimentary layer thickness obtained from microtremor measurements with the 'ground truth' borehole measurement is proposed for the purpose of seismic hazard reduction in urban areas that requires high-resolution near-surface geologic information. Traditionally many boreholes are required to obtain detailed information of sediment layers, which is time consuming and costly and in many cases not affordable for the research community. To reduce these constraints geostatistical data fusion of minimal borehole and more accessible microtremor surveys can be utilized to map the sedimentary layers at a high-resolution. Microtremor measurements can be conducted in urban areas whereas traditional methods of explosive and/or vibrating sources would not be acceptable. An existing collection of borehole data is available in Beijing area; nevertheless, it is still not enough to obtain comprehensive, high-resolution sedimentary thickness maps. In the summer of 2007 we had collected microtremor measurements over 600 sites in Beijing area. The inferred sedimentary thickness from the horizontal to vertical spectral ratio (H/V) of the microtremor data can be used to fill the vast gaps between isolated point measurements of sedimentary thickness provided by the sparser borehole data. Data fusion is accomplished by the geostatistical methods of Variograms, cross Variograms, and Cokriging in order to produce an estimated sediment layer map. Additionally, conditional simulation is utilized to enhance this estimation in order to provide an enhanced resolution sediment layer map for seismic hazard reduction in urban areas and other geophysics studies. This project is supported by the Ministry of Science and Technology of China with Project No. 2006DFA21650.

KEYWORDS: **sedimentary layer thickness, seismic hazard reduction, microtremor measurements, 'ground truth' borehole measurement, Geostatistical data fusion, Cokriging**

1. GEOSTATISTIC THEORY

According to Chilès (1999) Geostatistics attempts to quantify in a statistics sense spatial and/or time representations of natural variables. These variables historically are used to improve definitions of soil properties, rainfall, ore deposits, and atmospheric conditions. For this paper we are applying geostatistics to determine the Quaternary sediment layer depth in the Beijing metropolitan area in which the data is from borehole and microtremor studies (Chen et al 2008). Passive seismic sources such as ambient and man made sources called microtremors are processed by the horizontal to vertical spectral ratio (H/V) method to determine the depth of large velocity contrasts. These sources are received by a seismometer array which in turn can produce a large spatial and temporal data set that can be converted into determining sediment depth (Parolai et al, 2002) by examining the horizontal to vertical spectral ratio (Arai and Tokimatsu, 2004).

In this paper the borehole and microtremor data are fused together with the geostatistical method of cross-variogram; a new sediment depth map is constructed from the Cokriging. Cokriging is a geostatistical interpolation technique that allows one to better estimate spatially distributed values if the distribution of a secondary variate (microtremor data) has a larger number of samples with regards to the primary variate (borehole data). The cokriged data also has a standard deviation associated with it that can be used to provide additional confidence in the new data set. Cokriging is especially valuable if the primary variate is difficult or expensive to measure; by adding the secondary variate that is much easier and inexpensive to acquire cokriging can greatly improve interpolation estimates without having more primary variate samples.

A brief overview of the methods used in the paper is described first. Then, it is applied to the microtremor H/V and borehole data in Beijing area. The foundation of this approach is the variogram, which provides the spatial statistical relationship of the data; the output is interpolated to form the sediment thickness map via kriging or co-kriging (two data sets). The validity of the maps can be further enhanced through conditional simulation by reducing the spatial relationship semivariance.

1.1 Variogram

Variograms can attribute their roots to Fractal Theory according to Burroughs (1981) and are derived from the Correlogram (Dubin, 1998), which uses the covariance (requires the mean) rather than the semivariance. The Correlogram is shown in Eqn 1:

$$C(h) = E[Z(x) - m][Z(x+h) - m] \tag{1}$$

where E is the expected value, Z(z) the observation at x, m is the mean and h is the lag or separation distance. A variogram utilizes the difference between an observation Z(x) and a distance h away Z(x+h). The variogram is the backbone of this geostatistical method since it derives the spatial formula of semivariance of the observed data points with respect to how they are represented in space. The basic variogram is shown in Eqn. 2:

$$\gamma(h) = \frac{1}{2} Var[Z(x+h) - Z(x)] \tag{2}$$

where “Var” is the variance and note that in Eqn. 2 there is no dependence on the mean. The variogram over a data set as utilized for this paper is described by Eqn. 3:

$$\gamma(h) = \frac{1}{2N(h)} \sum [Z_i - Z_{i+h}]^2 \tag{3}$$

where zi is the observation sample at i, zi+h is the observation sample at i+h and N(h) is the total number of samples for lag interval h (Robertson, 1987).

1.2.1Variogram Models

The variograms can be defined to fit different models in order to define the semi-variance. These models are shown in Table 1, as well as graphically described in Figure 1.

Table 1 Different Variogram forms and with their associated equations

Form	Model
Linear	$\gamma(h) = C_0 + [h(C/A_0)]$
Spherical	$\gamma(h) = C_0 + C[1.5(h/A_0) - 0.5(h/A_0)^3]$ for $h \le A_0$ $\gamma(h) = C_0 + C$ for $h > A_0$
Exponential	$\gamma(h) = C_0 + C[1 - \exp(-h/A_0)]$
Gaussian or Hyperbolic	$\gamma(h) = C_0 + C[1 - \exp(-h^2/A_0^2)]$

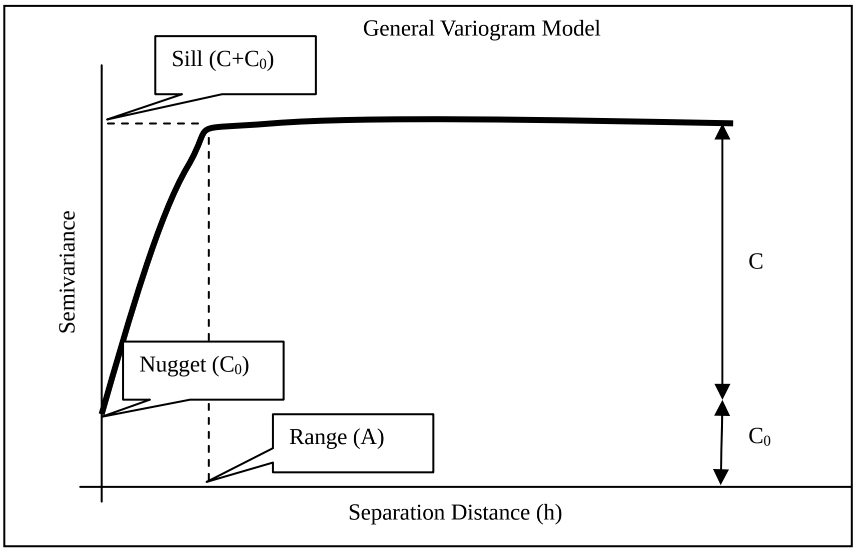

Figure 1 General Variogram Model

With respect to a variogram models as described in Table 1 and Figure 1 further definition is as follows.

The Nugget variance (Co) is the measurement error if non-zero or when h=0; the Sill (C + C0 = y) is when no further correlation between data points exists and in turn is equal to the variance of the set of observations (y); the range parameter (Ao) is the separation distance to the sill where there is no further spatial correlation; and C is the structural variance. Additionally, an isentropic variogram calculates semivariance based on distance in all directions around the observation. On the other hand, an anisentropic variogram includes a fixed direction in the range parameter when calculating semivariance since semivariance may be greater in a particular directions rather than in all directions. Another important factor is that correlation can be lost over a distance, which is termed drift. The models produced in this paper are based on an isentropic variogram model that attempt to minimize the nugget variance and drift.

1.2 Kriging

The variogram model defines the semivariances, which then can be utilized to form kriging weights. These are then utilized to interpolate between the known data to find the missing data in order to define a complete map of the spatial area of interest. The method utilized in this paper is called ordinary kriging and it estimates the unknown values by utilizing a linear combination of weights (ω) with the known data Z(xi) at the xi point Eqn. 4 where the sum of the weights (λi) equals one.

$$\hat{Z}(x_0) = \sum_{i=1}^{N} \lambda_i(x_0) Z(x_i) \tag{4}$$

Multiplying the covariances solves for the kriging weights as shown in Eqn 5;

$$\lambda_i = C^{-1} \cdot c_0 \tag{5}$$

where C is an (n+1)x(n+1) matrix and C(|h|=0) = Co + C with respect to Figure 1. The ordinary kriging weights are calculated in matrix form as Eqn 6;

$$\begin{bmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_N \\ \varphi \end{bmatrix} = \begin{bmatrix} \gamma(x_1,x_1) & . & . & . & \gamma(x_1,x_N) & 1 \\ . & . & & & . & . \\ . & & . & & . & . \\ . & & & . & . & . \\ \gamma(x_N,x_1) & & & & \gamma(x_N,x_N) & 1 \\ 1 & . & . & . & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(x_1,x_0) \\ . \\ . \\ . \\ \gamma(x_N,x_0) \\ 1 \end{bmatrix} \tag{6}$$

where φ is the Lagrange multiplier is used provided for the unbiasedness condition where the expected value of the estimated values equals the expected value of the observed values as shown in Eqn. 7

$$E\left[\hat{Z}(x)\right] = E\left[Z(x_0)\right] \tag{7}$$

Cokriging utilizes a cross-variogram model of two independent data sets. Cokriging is utilized when there is prohibitive cost or logistics in collecting actual observations when a second set of estimated observations can be used to fill in the data gaps. For this paper the primary data set is the borehole data that can be treated as the “ground truth” data and the secondary set, “guide data”, is the microtremor H/V data. This paper utilizes ordinary kriging as defined there is first and second order stationarity where there is similar statistical properties that describe the mean and the covariance of the observed data.

1.3 Conditional Simulation

Conditional Simulation is an iterative method that minimizes cross variance in an iteratively formed model developed by simulation of the observed data which improves the variogram model. For this let Z*(x) denote the kriging estimator of Z(x) on points x where Z(x) = Z*(x) + [Z(x) – Z*(x)] <= True value = Kriging Estimator + Kriging Error. Since Z(x) is unknown the kriging error is unknown and this can be initially solved through a non-conditional simulation. In a Gaussian case the error is independent from the “krig’ed” variable therefore mimicking the error by using a non-conditional simulation. The steps to develop a conditional simulation from a Gaussian data distribution are;

1. Create the non-conditional simulation (Zncs(x))
2. Krige Zncs(x) at its known data locations Zncs* (x)
3. Derive simulated error = simulated value – kriged value = Zncs(x) - Zncs(x)*
4. Add this error back to the kriging of the real variable Zcs(x) = Z*(x) + [Zncs(x) – Zncs* (x)] = Zncs(x) + [Z (x) – Zncs (x)] *
5. Keep doing this until the simulated error is reduced or minimized.

2. BEIJING MICROTREMOR STUDY

During the month of June 2007 a campaign of microtremor measurement survey was conducted for the assessment of local site effects of strong ground motion in the Beijing metropolitan area (Chen et al 2008). Over 600 sites with approximately 1 km spacing enclosed the area within Beijing's 5th beltway. The results of this study have shown the predominant resonant frequency, the ground motion amplification factor and the thickness of the uppermost soft sediment layer as shown in Figure 2a. The sediment layer thickness agreed with sparse borehole data and further enhanced the depth resolution of previous state of the art studies that mapped out the sediment depths under Beijing such as Ding et al (2004) as shown in Figure 2b.

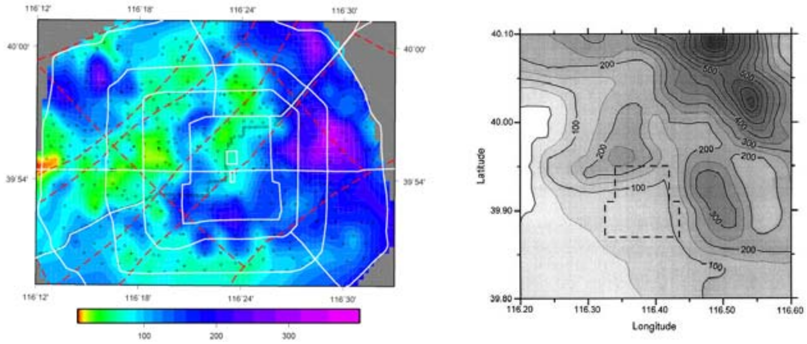

Figure 2. (a) Microtremor sediment depth study results (chen et al, 2008) with major roadways (white lines) and fault lines (red dashed lines); and (b) Quaternary depth layer map (Ding et al, 2004)

Geostatistical methods were utilized to further enhance the microtremor sediment layer depth map as shown in Figure 2. Since borehole data is costly and microtremor data is relatively inexpensive this paper will utilized the Geostatistical processes previously described to fuse these two data sets together.

Utilizing the process as shown in Figure 3 where borehole (ground truth data) is combined with the microtremor data in order to develop a cokriged sediment layer depth map. Additionally this map will have an estimate of spatial standard deviation that will provide a statistical sense to its validity. Further enhancement to this map may be accomplished through conditional simulation.

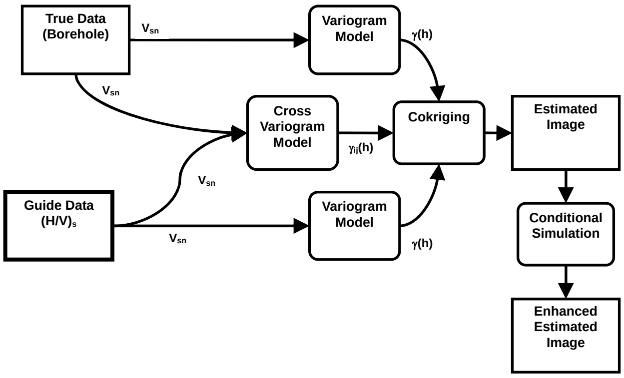

Figure 3 Geostatistical Process used in this paper

The data fusion through co-kriging uses the ground truth data consisting of 31 borehole studies and 585 microtremor H/V measurements. An Exponential model was chosen for the primary variate, borehole, variogram with the following parameters; nugget (Co) = 0.0010, sill (Co+C) = 0.5200, range (Ao) = 0.10 and h=0.13. Whereas the microtremor variate utilized a spherical model with the following parameters; Co=0.1320, Co+C=0.3510, Ao=0.07, and h=0.13. Fusing these data resulted in a cross variate exponential model with the following parameters; Co=0.0010, Co+C=0.4940, Ao=0.12 and a lag distance (h) = 0.13m. The Cokriging Cross Validation yielded a good regression coefficient of 0.935, which implies that the variates are highly correlated.

Figure 4a is a kriged version of the borehole data only. Since the boreholes are loosely spaced and are some distance from each other the standard deviation, Figure 4b, of the data set increases away from each borehole sites, and increases significantly at the other edges.

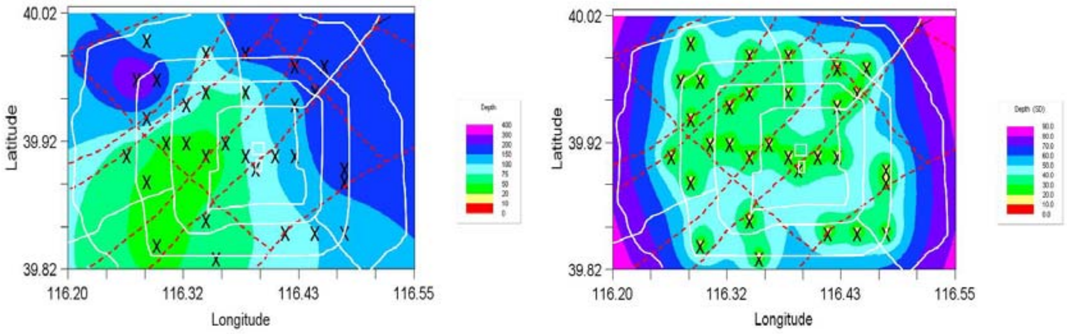

Figure 4 (a) Kriged results of only the borehole data; and (b) its standard deviation with major roadways (white lines), fault lines (red dashed lines), and X's represent the borehole sites

The microtremor data is included to form the cokriged sediment depth layer map as shown in Figure 5a. Since the additional microtremor data is included the sediment depth map has a higher resolution then the original borehole data only sediment depth map as shown in Figure 4a. Additionally, the standard deviation of the data is minimized in the cokriged version as seen in Figure 5b. With this sediment map we could present its results with a degree of confidence that the depth error is within 30 to 40 meters in the majority of this sediment depth representation of the Beijing metropolitan area.

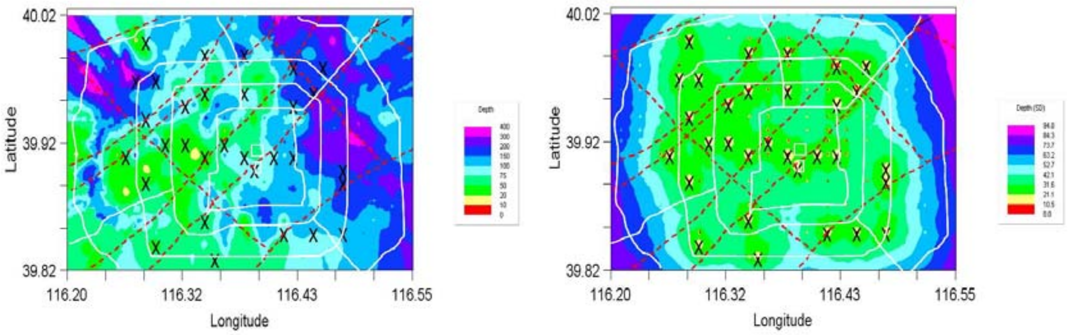

Figure 5 (a) Cokriged results; and (b) its standard deviation with major roadways (white lines) , fault lines (red dashed lines), and X's represent the borehole sites

Utilizing conditional simulation of 1000 simulation sets, an arbitrary number, the semivariance was minimized and produced the final sediment layer map of the Beijing metropolitan area as shown by Figure 6a. Of interest is the standard deviation, Figure 6b, of the deeper holes and some of the internal borehole sites are higher than the outer edges, which is somewhat contrary to the previous cokriged data set. Additionally the data set resolution has been reduced or smoothed, which increases as the number of conditional simulations increase.

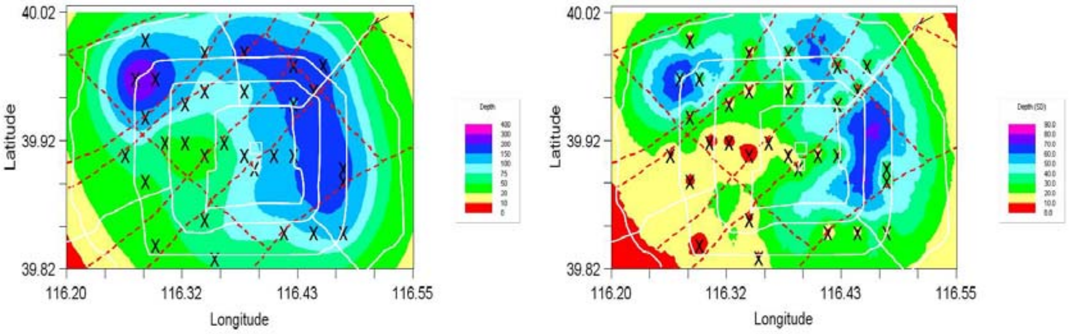

Figure 6 (a) Conditional Simulation Cokriged results; and (b) standard deviation with major roadways (white lines) , fault lines (red dashed lines), and X's represent the borehole sites

3. DISCUSSION

As shown in the above analysis, geostatistical methods can increase the resolution of the borehole and microtremor studies when combined through cokriging. Additionally it provides a sense of how the data is represented in the unknown areas, where no depth information exists, in the form of the standard deviation. The cokriged data set shown in Figure 5a is a good generalization of the soft sediment thickness of Beijing area, which reduces the semivariance as shown by the standard deviation of the data as compared to the borehole only data. With conditional simulation it improved the standard deviation of the layer depth in the south-west quadrant of the Beijing metropolitan area, but did not further improve other portions in the area.

4. CONCLUSION

This paper demonstrates a method to increase the resolution of soft sediment layer depth estimate through the data fusing of borehole and microtremor data while providing a confidence level of the validity of the produced data. These data may provide additional detail to geotechnical engineers for hazard reduction mitigation in the Beijing metropolitan area. Future work should examine methods to weight the ground truth (borehole data) and the soft data (microtremors) in a more rigorous way to explore the possibility to generate a sediment thickness map with even higher resolution.

REFERENCES

Arai, H. and Tokimatsu, K. (2004). S-Wave Profiling by Inversion of Microtremor H/V Spectrum. *Bulletin of Seismological Society of America* **94:1**, 53-63.

Burrough, P.A. (1981). Fractal dimensions of landscapes and other environmental data. *Nature* **294**, 240-242

Chen, Q., Liu, L., Wang, W. and E. Rohrbach, E. (2008). Site effects on earthquake ground motion based on microtremor measurements for metropolitan Beijing. *Chinese Science Bulletin* (in press)

Chilès, J-P and Delfiner, P. (1999). Geostatistics Modeling Spatial Uncertainty, Wiley-Interscience, New York, US

Ding, Z., Romanelli, F., Chen, Y.T., and G.F. (2004). Realistic Modeling of Seismic Wave Ground Motion in Beijing City. *Pure appl. geophys* **161**, 1093–1106

Dubin, R.A. (1998). Spatial Autocorrelation: A Primer. *Journal of Housing Economics* **7:4**, 287-377

GS+ Geostatistics for the Environmental Sciences, Gamma Design Software, Plainwell Michigan

Parolai, S., Bormann, P. and Milkereit, C., (2002). New Relationships between Vs, Thickness of Sediments, and Resonance Frequency Calculated by the H/V Ration of Seismic Noise for the Cologne Area (Germany). *Bulletin of the Seismological Society of America* **92:6**, 2521-2527

Robertson, G.P. (1987). Geostatistics in Ecology: Interpolating with Known Variance. *Ecology* **68:3**, 744-748