Hydraulic Engineering
M. A. Ports (ed)
ASCE 1987

BORDEN FIELD DATA AND MULTIVARIATE GEOSTATISTICS

Donald E. Myers[1]

Abstract

A large-scale field experiment on natural gradient transport of solutes in groundwater was conducted in an unconfined sand aquifer in Borden, Ontario by extensive monitoring after the pulse injection of two inorganic tracers and five halogenated organic chemicals. The data is presented as a case study in modeling variograms and cross-variograms including the case where some variables are under-sampled. Co-kriging allows the interpolation and hence contouring incorporating both spatial and inter-variable correlation. Numerical results are presented comparing results obtained by other methods and co-kriging.

Introduction

A large-scale field experiment on natural gradient transport of solutes in groundwater was conducted over a two year period near Borden, Ontario (Roberts and Mackay, 1986). Known masses of two inorganic tracers and 5 halogenated organic chemicals were injected and monitored at 5000 sites. Samples were collected (nearly) daily and at several depths.

Because samples are indexed both with respect to location and date, the data set was first re-organized into two categories of data files; one for each date including all locations, one for each location including all dates. In the case of the latter many of the locations far from the injection wells did not show detection of the solutes at all or only for a small number of dates. There could be at least two general objectives in applying multi- variate geostatistics to data sets such as the Borden data; first, characterize and model spatial correlation
for each solute, characterize and model spatial correlation for each pair of solutes as well as the temporal correlation and cross-correlation, secondly, utilize these characterizations for kriging and/or cokriging to contour the plumes of the solutes. After computing and plotting of variograms, cross-variograms for all dates and pairs of variables; results for two dates and two solutes were selected for presentation here.

Theory

A complete presentation of the theory is given in Myers (1982, 1988). Denote by $Z_j(x)$ the value of solute j at location x, then the spatial (cross) correlation with $Z_k(x+h)$ is quantified by the (cross) variogram

$$\tau_{jk}(h)=\mathrm{Cov}\{Z_j(x+h)-Z_j(x),Z_k(x+h)-Z_k(x)\} \qquad (1)$$

For m solutes or variables, let $Z(x)=[Z_1(x),...,Z_m(x)]$ then the general co-kriging estimator

[1]Professor of Mathematics, University of Arizona, Tucson, Arizona 85721

is given by

$$Z^*(x)=\Sigma Z(x_i)\Gamma_i, \text{ where the } \Gamma_i \text{ are weight matrices.}$$

The variograms/cross-variograms must satisfy certain positive definite conditions as described in Myers (1988). While it is appropriate to model the variograms directly, it is more convenient to model the cross-variograms indirectly as indicated in Myers (1982,1988), namely by the use of the variograms of the sums and differences. Because variogram, cross-variogram modeling involves some subjectivity it is also desirable to validate the models against the data recognizing the robustness and continuity of the kriging, co-kriging estimators with respect to the models. This is discussed in Myers (1985). The co-kriging estimator can be used in either the full-sampled or the undersampled form, as shown in Carr, Myers and Glass (1985) the program can accomodate either form and as noted in Myers (1988) it makes little difference whether estimating all variables or only one.

Numerical Results

For an initial evaluation of the application of geostatistics to this data set. Two dates were considered 9/8/82 for solutes chloride, bromide, and 10/5/82 for solutes chloride and carbon tetrachloride. At this point the selection was made simply because the variogram plots appeared more amenable. Plots of the variograms for chloride (both dates) and for Carbon Tetrachloride as well as the cross-variogram for Chloride vs Carbon Tetrachloride are shown in the figures.

Because of the substantial differences in the magnitudes of the values for the two solutes for both dates which made discerning the plots more difficult the data was re-scaled but this re-scaling was removed after modeling prior to co-kriging. The variograms and cross-variograms were cross-validated in two stages; first variograms were cross-validated using kriging then each triple (vario- grams for each of the two solutes together with the cross variogram) was cross validated using co-kriging. The co-kriging program is an up-dated version of the one given in Carr, Myers and Glass (1985) but which is written in Microsoft FORTRAN 77 and hence useable on either a micro or a VAX. The cross-validation included the use of the scatter plots and correlations between estimated, data values at sample locations as well as those for estimated vs estimation errors together with mean squared error and mean kriging variances. The particular choices of the solutes also illustrated the effect of the correlation between the solutes on the co-kriging; for chloride and carbon tetrachloride the sample correlation is 0.489 whereas for chloride and bromide it was 0.983.

The cross validation results are given in the following tables

TABLE 1
Cross Validation for Chloride and Carbon Tetrachloride

	Chloride	CTET
Sample locations	313	241
Mean error	2.43	-0.076
Mean Square error	21334	25.66
Stand. Mean Square error	1.002	1.096
Correlation est. vs act.	.82	.82
Correlation est. vs error	.067	.007

Table 2
Cross validation for Chloride and Bromide

	Chloride	Bromide
Sample locations	339	339
Mean erro	-0.037	-0.49
Mean Square error	3530	3358
Stand. Mean Square error	1.10	1.07
Correlation est. vs act.	.76	.75
Correlation est. vs error	.024	.009

Following the modeling of the variograms and cross variograms both solutes were kriged (separately) and co-kriged (jointly) on a grid to compare the two techniques. As might be expected the results from co-kriging were not effected by reducing the number of sample locations for bromide or carbon tetrachloride while using the full data set for chloride.

Acknowledgements

Although the research described in this article has been funded wholly or in part by the U.S. Environmental Protection Agency through a Cooperative Research Agreement with the University of Arizona, it has not been subjected to Agency review and therefore does not reflect the views of the Agency and therefore no official endorsement should be inferred.

The author thanks D.C. Bouchard, US-EPA, Ada, OK for providing access to the data and a copy of the technical report. Computations were provided by Zhang Renduo.

References

Carr,J., Myers,D.E. and Glass,C., "Co-kriging:A Computer Program," Computers and Geosciences, 11 (1985) 111-127

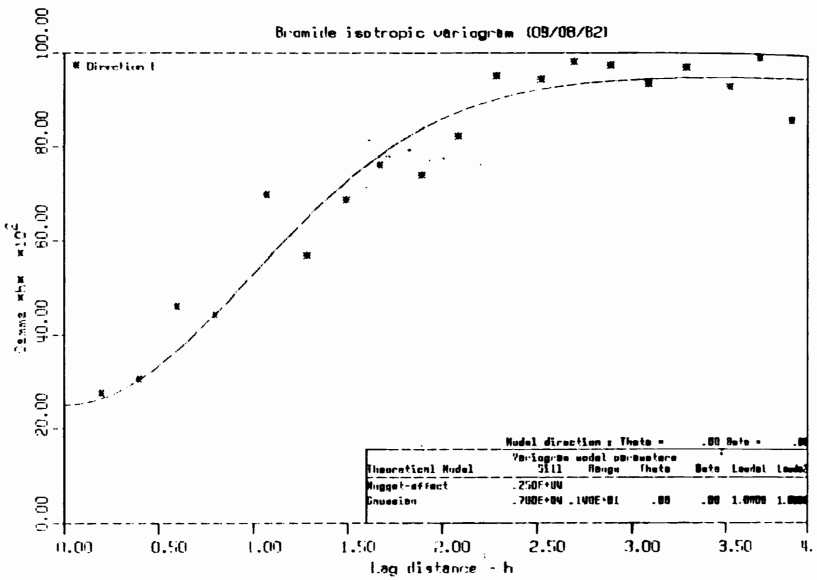
Myers,D.E., "Matrix Formulation of Cokriging," Math. Geology, 14 (1982) 249-257

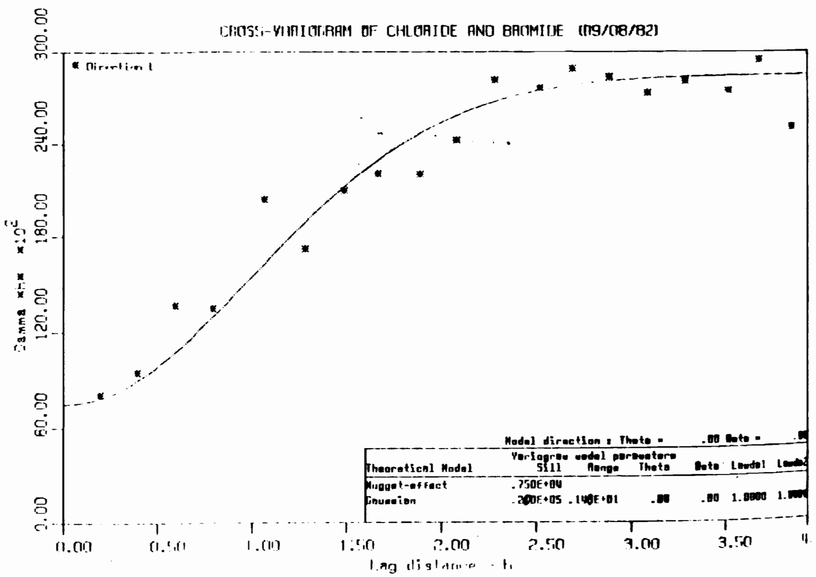
Myers,D.E., "On Some Aspects of Robustness," Sciences de la Terre,24 (1985) 63-79

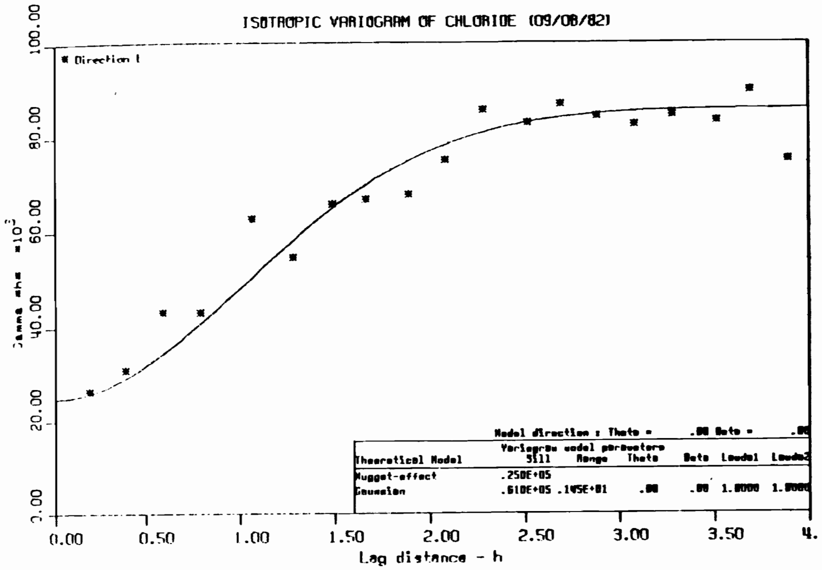
Myers,D.E., "Interpolation with Positive Definite Functions," Sciences de la Terre,28 (1988) 251-265

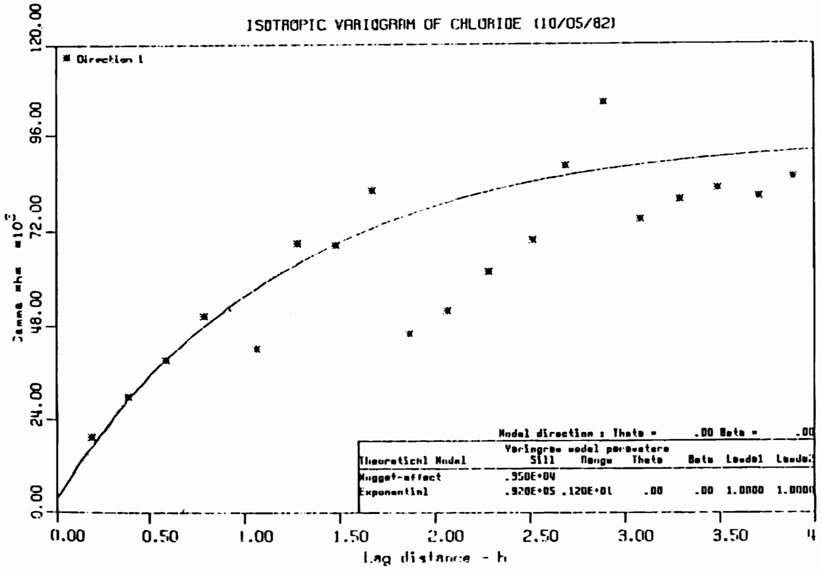
Myers,D.E., "Multivariate Geostatistics for Environmental Monitoring," Sciences de la Terre,27 (1988) 411-427

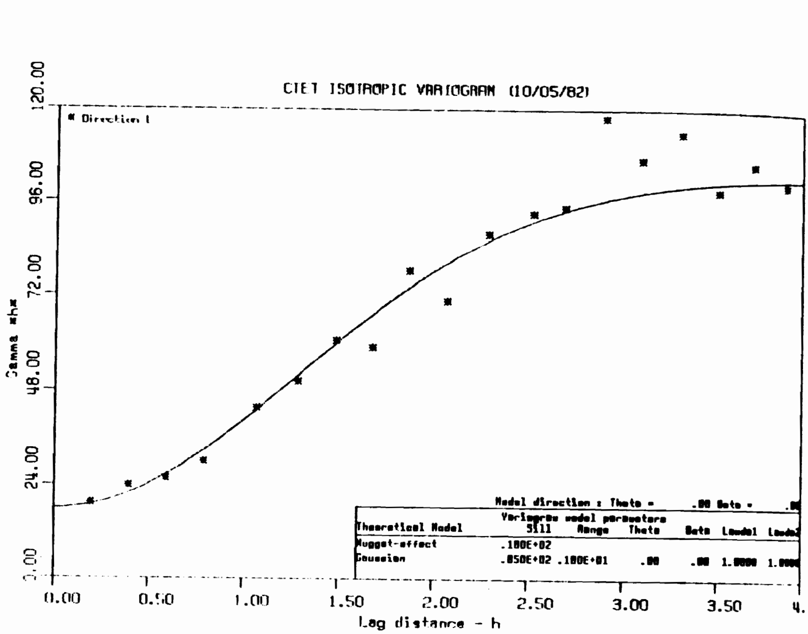
Roberts,P.V. and MacKay,D.M.(eds), A Natural Grad- ient Experiment on Solute Transport in a Sand Aquifer. Tech.Report 292, Dept. Civil Eng., Stanford Univ.

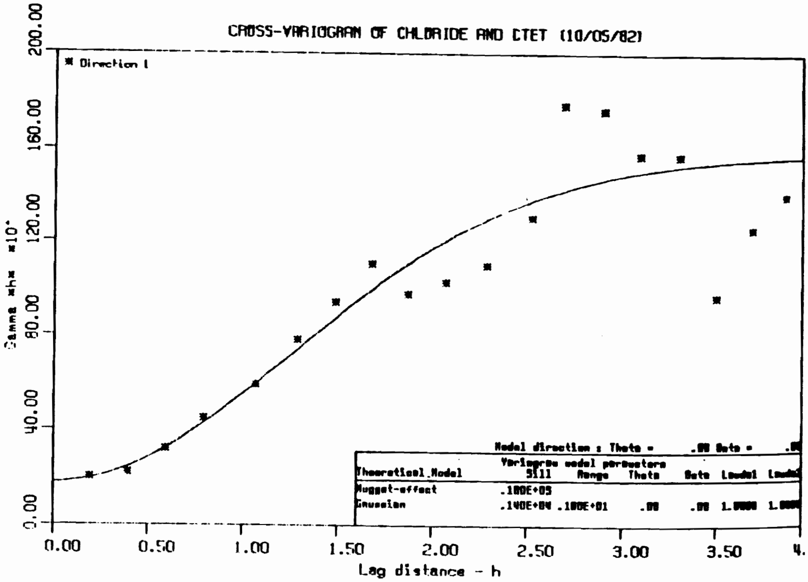

Bromide isotropic variogram (09/08/82)
* Direction 1
Gamma *h* *10^2
0.00
20.00
40.00
60.00
80.00
100.00
0.00 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4.
Lag distance - h
Model direction : Theta = .00 Beta = .0
Variogram model parameters
Theoretical Model Sill Range Theta Beta Lawdal Lawda2
Nugget-effect .250E+04
Gaussian .700E+04 .140E+01 .00 .00 1.0000 1.0


CROSS-VARIOGRAM OF CHLORIDE AND BROMIDE (09/08/82)
* Direction 1
Gamma *h* *10^2
0.00
60.00
120.00
180.00
240.00
300.00
0.00 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4
Lag distance - h
Model direction : Theta = .00 Beta = .0
Variogram model parameters
Theoretical Model Sill Range Theta Beta Lawdal Lawda2
Nugget-effect .750E+04
Gaussian .2?0E+05 .14?E+01 .00 .00 1.0000 1.0


ISOTROPIC VARIOGRAM OF CHLORIDE (09/08/82)
* Direction 1
Gamma *h* *10^3
0.00
20.00
40.00
60.00
80.00
100.00
0.00 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4.
Lag distance - h
Model direction : Theta = .00 Beta = .00
Variogram model parameters
Theoretical Model Sill Range Theta Beta Lawdal Lawda2
Nugget-effect .250E+05
Gaussian .610E+05 .145E+01 .00 .00 1.0000 1.0000


ISOTROPIC VARIOGRAM OF CHLORIDE (10/05/82)
* Direction 1
Gamma *h* *10^3
0.00
24.00
48.00
72.00
96.00
120.00
0.00 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4
Lag distance - h
Model direction : Theta = .00 Beta = .00
Variogram model parameters
Theoretical Model Sill Range Theta Beta Lawdal Lawda2
Nugget-effect .950E+04
Exponential .920E+05 .120E+01 .00 .00 1.0000 1.0000


CTET ISOTROPIC VARIOGRAM (10/05/82)
* Direction 1
Gamma *h*
0.00
24.00
48.00
72.00
96.00
120.00
Model direction : Theta = .00 Beta =
Variogram model parameters
Theoretical Model
Sill
Range
Theta
Beta
Nugget-effect
.100E+02
Gaussian
.850E+02 .100E+01
.00
.00
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.
Lag distance - h


CROSS-VARIOGRAM OF CHLORIDE AND CTET (10/05/82)
* Direction 1
Gamma *h* *10'
0.00
40.00
80.00
120.00
160.00
200.00
Model direction : Theta = .00 Beta =
Variogram model parameters
Theoretical Model
Sill
Range
Theta
Beta
Nugget-effect
.100E+05
Gaussian
.140E+04 .100E+01
.00
.00
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.
Lag distance - h