

Clark, W.V. Harper and K.L. Basinger, "*MUCK -- A novel approach to co-kriging*" in Proc. DOE/AECL '87 Conference on Geostatistical, Sensitivity and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modelling, September

Proceedings of the Conference on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modeling

San Francisco, California, September 15-17, 1987

Sponsored by U.S. Department of Energy and
Atomic Energy of Canada Limited

Edited by: Bruce E. Buxton

Battelle Memorial Institute, Columbus Division, Columbus, Ohio

BATTELLE PRESS

MUCK - A Novel Approach to Co-Kriging

ISOBEL CLARK, Geostokos Ltd., London, England

KAREN L. BASINGER and WILLIAM V. HARPER, Battelle Office of Nuclear Waste Isolation, Columbus, Ohio

Abstract

Multivariate geostatistical estimation – “co-kriging” – has been used in practice for well over a decade. The technique as developed by Matheron has been documented in many case studies.

There are disadvantages to the standard approach, however, not the least of which is that the variables which are measured must be sited at the same geographical location. For example, it is necessary that if an exploration project is considering several minerals, they will tend to measure all of them on the same samples.

There are cases in which this type of information is not available. Indeed, the motivation for using a multivariate approach may be to specifically improve estimation in areas where few variables have been measured.

A variation in co-kriging will be presented which does not place this restriction on the modeling and estimation processes. A case study will be given which includes the added complication of strong trend. Hence the acronym for this new method - Multivariate Universal Co-Kriging.

Introduction

The term “co-kriging” appears to have been coined by Matheron himself in his original work, *The Theory of Regionalized Variables and Its Applications* (1971). A review of this original approach is useful before starting to generalize.

It is common in geological sampling to measure more than one variable at each sample location. For mining purposes, for example, many mineral grades might be measured, plus perhaps the width of the lithological unit. In hydrology, one might possibly measure porosity and permeability, hydraulic head, grain size, and so on. Some of these variables might be correlated. Some of them may not have been measured at all locations. If the value of a variable at an unsampled location is to be estimated, the spatial structure (semivariogram) for that variable must be established. The sample values may then be used to produce a “kriging” estimator of the variable at that location.

Suppose, however, that this variable has been fairly sparsely sampled; it may be more expensive to obtain values, or perhaps just very difficult. Suppose, in addition, that this variable is significantly correlated with one or more of the other variables that have been measured. Matheron studied the question as to whether these correlated variables could be used to estimate the variable of particular interest. He developed the concept of a “cross semivariogram” between two variables and, from this structure, a set of kriging equations, which he dubbed “co-kriging.” The details of this theory are thoroughly documented by Matheron (1971) and in numerous papers, notably those by Guarascio (1976) and Myers (1982, 1983, 1984), so that only those points that are relevant to the present discussion need be discussed here.

When Matheron developed co-kriging, he supposed that there would be (effectively) a single sampling campaign. Some or all of the variables of interest would be measured at each location. The “cross semivariogram” was defined as follows:

$$\gamma_{gf}(h) = 0.5 E[(g_i - g_j)(f_i - f_j)]$$

where g_i, g_j are the measured values of variable g at locations i and j

where f_i, f_j are the measured values of variable f at locations i and j

and h is the distance between the two samples.

Under the usual intrinsic hypothesis and the assumption of no significant drift, the estimation of the value g_i at an unsampled location gives rise to a set of co-kriging equations. The solution of these equations produces a minimum variance unbiased estimator in the form of a linear combination of the neighboring values of both variables. This set of equations can be generalized with reasonable ease to the case where a significant drift is present (Myers, 1982).

Although co-kriging does not seem to be in routine use in geological applications, well-documented case studies illustrate the use of the technique in various situations (cf. Carr and McCallister, 1985; and Unal and Haycocks, 1986).

Some problems associated with co-kriging

The theory of co-kriging seems to be well founded. However, it has been noted by various authors (Carr et al., 1985) that its application in practice gives rise to some real difficulties. These are mostly encountered in the calculation and modelling of the cross semivariogram graph. Myers (1982) suggests that difficulties in semivariogram modelling can be reduced by using a pseudo-cross semivariogram comprising the *sum* of the values squared (rather than just the difference). This graph is modelled and then “adjusted” by subtracting the single-variable semivariogram models and dividing by two.

Myers (1982) states that “it is simpler to utilize cross variograms than to use cross covariances, but this is not always possible”. He formulates his approach to co-kriging, therefore, almost completely with cross covariance functions and generalizes to the semivariogram function. These calculations require stationarity of both mean and variance rather than the less restrictive intrinsic hypothesis. He also states that the cross covariance function and the single-variable covariance functions must obey the Cauchy-Schwartz inequality in the same way that a normal covariance must be less than the product of the two standard deviations. At first sight, this seems sensible, but it is difficult to show mathematically other than at the origin of the graph. Myers also asserts that the Cauchy-Schwartz inequality should hold between the cross semivariogram and the univariate semi-variograms. There is no justification for this statement.

In fact, given the formulation of the traditional cross semivariogram function, there does not seem to be any practical reason why it should be nonnegative. If the two variables under consideration are inversely correlated at a significant level, then the cross semivariogram function would have a stable form but would show negative values for $\gamma(h)$. This means that we have to model a “semi-variance” graph that generally decreases as the distance increases. This could lead to some very strange kriging weightings for distant samples. Negative kriging variances might also be encountered. If the variables are effectively uncorrelated, the graph will oscillate around zero - a very unsatisfactory behavior from the point of view of building a stable set of kriging equations.

The problems arising in the usage of the cross semivariogram graph might be due to the original formulation of the function, rather than to instabilities in the modelling or the sample configuration. The choice of the traditional form follows naturally when the derivation is made first for the covariance function and only then generalized to the semivariogram graph. It can easily be shown that the above semivariogram definition is obtained by subtracting the cross covariance function from the normal statistical covariance at distance zero. The authors propose the reverse procedure, deriving the cross semivariogram function from the concept of the single variable semivariogram function and the intrinsic hypothesis.

An Alternative Formulation

The semivariogram for, e.g. variable g , is usually defined as

$$\gamma_g(h) = 0.5 E[(g_i - g_j)^2]$$

using the same notation as before. This definition is justifiable on many grounds, most of them mathematical. Intuitively, this approach may be justified as follows: the ultimate goal is to

produce an estimate for an unsampled location that will be as close as possible to the actual value; the difference between the estimate and the actual value is the “estimation error”; the best estimator will be one with the smallest errors – in mathematical terms, mean zero and minimum variance. In short, a kriging system seeks to minimize the expected value of the squared errors, i.e., the squared difference between the known sampled values and the unknown unsampled value. Therefore, it seems only sensible to build a continuity or “predictability” model to describe the squared difference between values at different locations.

For variable f the equivalent form is

$$\gamma_f(h) = 0.5 E[(f_i - f_j)^2]$$

It is suggested that the cross semivariogram be defined as

$$\gamma_{gf}(h) = 0.5 E[(g_i - f_j)^2]$$

This formulation follows the intuitive logic described for the ordinary semivariograms. The only real difference between the single-variable semivariogram and the two-variable cross semivariogram is that the values at different sample locations will be values of different variables.

A possible difficulty in the practical application of either type of cross semivariogram will occur if the two variables are of widely differing scales. This problem is analogous to that encountered in calculating classical covariance between such variables, where precision may be lost due to the numerical values being of different orders of magnitude. A suggested solution is to “standardize” the individual variables to give each an average of zero and a sill on the univariate semivariogram of unity. The situation is further complicated in the presence of trend and is suggested as a topic for intensive study at a later date.

The formulation of the cross semivariogram function shown above has some advantages over the classical Matheron method. First, and perhaps most important, it cannot become negative. If the samples are highly negatively correlated, the graph will rise precipitately rather than fall below zero. If the variables are uncorrelated, the differences will show as a constant “nugget effect” – a flat horizontal line on the graph. This function, then, may be modelled in the same way as a normal single variable semivariogram with no need to introduce intermediate graphs.

It does not appear easy to establish a relationship between this type of cross semivariogram and the cross covariance function. The mathematics is not particularly complex, but it involves the knowledge of a stationary mean value for each of the two variables. If a covariance exists, i.e., if the variables are stationary over the study area, then of course the mean will be stationary. However, the construction of the cross semivariogram function does not require that the mean be stationary and may, therefore, be more geologically valid.

Our formulation of the cross semivariogram has been expressed as though there were no significant drift or trend over the study area. This is not really a restriction. If there is a significant trend, then it must be removed in some way before any of the semivariograms can be calculated. In that case, then all semivariograms are calculated as the difference between the residual values at each sample location.

The major advantage

It might appear that the reformulation of the cross semivariogram function is justified on mathematical grounds alone. The removal of the extra step in the modelling is an unexpected but welcome bonus. However, it must be said that neither of these features is the main motivation for the development of a new approach to cokriging.

To explain the reasons for the current development, the motivation for the original development of co-kriging must be examined. The major application for co-kriging over the last two decades or so has been to estimate one particular variable that has been undersampled compared to the others. For example, it is relatively easy to measure base-metal content in a massive sulfide deposit. It is, however, costly and time consuming to measure the specific gravity of every drill-core section. On the other hand, it is reasonably easy to establish a correlation between metal value and specific gravity. Co-kriging is one way in which the density might be estimated more accurately without recourse to vastly increased sampling costs. The other major application (Myers, 1982) appears to be in estimating linear combinations of various parameters at one time.

There is a fundamental difference between these requirements and those of the program that initiated the present research. The study concerned the potentiometric levels within various aquifers in a particular study area with a view to ultimately predicting fluid flow and travel path parameters. Because of the nature of the sampling available, those boreholes having reliable observations of the potentiometric pressure in one aquifer tended to be at different locations from the boreholes used to sample another aquifer. The sample data, then, could not be used to calculate a semivariogram such as that described in the classical geostatistical literature because these require multiple measurements at a significant proportion of the sample locations.

Our reformulation of the cross semivariogram enables all potential pairs of sample values to be used, including values at the same location should they be available. This feature is invaluable in helping to establish the nugget effect, which must be present in a cross semivariogram. Both a more stable shape for the graph (because it is based on more pairs) and a more efficient estimate for all parameters required for the model-building stage are obtained. Standard semivariogram models may be used without problems, and they have been found to suit all of the cross semivariograms encountered so far.

Co-kriging

If the proposed form of the semivariogram is acceptable and the underlying intrinsic hypothesis that “difference in value” is a function only of distance between sample locations, some kind of optimal estimator can be developed. Only the linear estimator with or without significant trend will be considered in this paper, i.e., those methods known as “ordinary kriging and universal kriging.” These techniques are based on the assumption that a linear combination of the sample values provides a sufficient estimator for an unsampled location, and that a good measure of the accuracy of this estimator is the variance of the error of estimation thus produced. The kriging equations are a direct result of minimizing this estimation variance under the constraint that the estimator be unbiased.

Carrying out such a derivation on the basis of the new cross semivariogram model presents little difficulty other than that of algebraic complexity. The resulting set of kriging equations are identical to those traditionally used for co-kriging. Myers (1982) formulated the co-kriging

system as a set of matrices that is invaluable for clarifying the situation when all measurements are made at all locations. It becomes a little unwieldy when there are a large number of “missing” or undersampled variables, because rows and columns must be deleted from each matrix. In this situation, where very few of the locations can be assumed to have more than one variable measured, the matrix approach is of limited use. The equations have, therefore, been expressed in an algebraic notation and then reformulated into matrices for the purposes of computer programming.

A brief appendix is attached that provides the algebraic expressions for the co-kriging equations when estimating the value of one variable at an unsampled location. A short discussion is given here.

In formulating the layout for the co-kriging equations, the following information is required:

1. The location of the sample (geographical coordinates).
2. The variable measured at this location.
3. The value observed for this variable.

It is obvious from the above that if more than one variable *has* been measured at a single location, it should be treated as more than one sample in the data set. The bulk of the co-kriging equations then look identical to an ordinary kriging system:

1. The left-hand matrix will contain semivariogram values for all possible combinations of sample values.
2. The right-hand “vector” will contain the semivariogram values between each sample and the unsampled location being estimated.
3. The left-hand “vector” will contain the weighting factors allocated to each sample to produce the optimal estimation of the unknown value.

There are two major differences between this and the usual kriging system of equations. First, one must carefully keep track of which variable is measured on each sample so that semivariogram values from the correct model will be calculated, whether single-variable or cross semivariogram. This presents few problems. Second, one must apply the usual constraints to ensure that the estimator remains unbiased. In the ordinary kriging system, the weighting factors are constrained to sum to unity, thus ensuring unbiasedness in the absence of trend. This adds one equation to the system and one “unknown” to the vector of weighting factors. In universal kriging, a constraint is added for each mathematical term in the expression of the trend, so that the estimator remains unbiased. If the trend contains six terms, six equations will be added to the kriging system and six “unknowns” to the final solution.

In co-kriging, similar constraints must be used to ensure that the estimator is unbiased.

Considering the case without trend, previous authors (op. cit.) have suggested the following conditions:

1. Identify which samples have been measured for the same variable as that being estimated. The sum of the weighting factors for these samples must be unity.
2. For *each* other variable, identify the samples for which this variable has been measured. The sum of the weighting factors for these samples must be zero. This process is repeated for each variable encountered in this sample set.

These conditions are extremely restrictive. There is no reason, for example, why the individual variables should all have factors summing to zero. If the study area does exhibit stationarity and the mean value for each variable could be estimated, the constraints could follow any number of different patterns. However, this question is outside the current scope of this paper. In addition, adoption of the above constraints leads to some great simplifications when computer programming.

In short, then, to ensure that the co-kriging estimator is unbiased, there must be as many constraints as there are observed variables in the sample set, or at least in the set being used for estimation. In the case where significant trend exists, universal co-kriging constraints must be added for each term in the (polynomial) trend that has been identified for each variable. For example, if there are three variables with no trend on the first, three terms in the trend for the second, and six terms for the third, then $1 + 3 + 6 = 10$ equations must be added to the co-kriging system. That is, ten extra rows and columns to the left-hand matrix, and ten items to each of the right-and left-hand vectors.

Considering ordinary kriging without drift as a special case of universal kriging with one term (constant) in the trend, the process of co-kriging might be summarised as follows. A set of equations is developed that is remarkably similar to those of single-variable kriging, except that the constraints on the weighting factors require the addition of enough equations to cover all of the terms in all of the trends for all of the variables being used in the estimation. Otherwise, the only noticeable difference is that there are more sample values available in making each estimate.

Multivariate co-kriging

In his 1982 paper, Myers pointed out that the estimation of one (possibly undersampled) variable was not really the prime purpose of co-kriging. When expressed in matrix form, it is clearly seen that some or all of the variables may be estimated at the same time without a significant increase in computing time or cost. This is simply because the “left-hand matrix” of intersample semivariogram values will remain the same no matter which variable is actually being estimated at the unsampled location. That is, values can be estimated for all of the variables at a single location and the matrix on the left-hand side would be the same.

The two vectors involved in the co-kriging system will change, of course. The vector on the right-hand side contains the semivariograms between each sample and the value at the unsampled location. This will obviously change if a different variable is being estimated. The variable under estimation will also affect the constraints on the weighting factors. If estimated values are required for several variables at the unsampled location, there will be as many right-hand-side vectors as there are variables. In direct consequence, there will be as many solution vectors as there are variables. However, since the left-hand side remains the same, the solution for all of the vectors may be carried out simultaneously. The primary implication of this is that many variables may be estimated for the cost of one.

Multivariate universal co-kriging

Co-kriging may be used to estimate an inadequately sampled single variable, or it may be used to estimate several variables simultaneously. In the latter case, the estimation of each variable is improved, in that more information may be used, and the cost is scarcely greater than estimating any one of the variables singly. To avoid confusion as to the purpose of this estimation process,

it is proposed that the term “multivariate” be used to emphasize that more than one variable is being estimated. Because the technique is also applicable in the presence of significant trend, the traditional term “universal” should be included; thus the term “multivariate universal co-kriging” or MUCK, is derived.

In summary, then, a new definition of the semivariogram graph is proposed which meshes intuitively with the concept of the original semivariogram in use for single variables. The inclusion of this new form results in a system of equations remarkably similar to that obtained using the old form. This system of equations is used to obtain “optimal” estimates for all of the variables at each unsampled location simultaneously. This approach is referred to as multivariate universal co-kriging (MUCK).

A Case Study - The Pennsylvanian and Wolfcamp Aquifers

The area of study is a region in the northern portion of Texas in the Panhandle region (Figure 1).

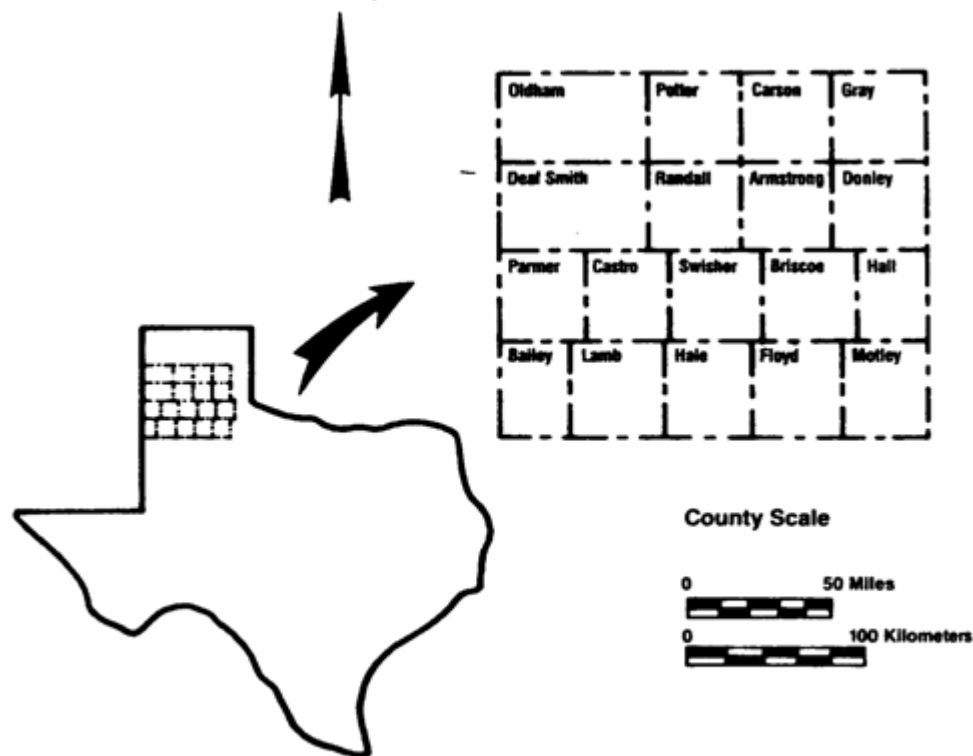


Figure 1. Location map for the geostatistical analysis.

The sample data are potentiometric values from deep brine aquifers called the Pennsylvanian and the Wolfcamp, which in some areas are almost 8,000 feet below the ground surface. The variable considered in this study is the potentiometric or piezometric level, expressed as the elevation (in feet) above sea level of water in a well at that location.

Deaf Smith County in northern Texas is a possible site for a future high-level nuclear waste repository. The Pennsylvanian and Wolfcamp aquifers underlie the salt bed, which would be the host for such a repository (cf. Bair et al., 1985). The analysis of these aquifers is a vital

component in the characterization of this site, because a breach of the repository could cause hazardous material to be released into the aquifers. One of the requisite parameters in determining flow rates and directions is the potentiometric level, because fluids tend to flow from areas of higher potentiometric values to those that are lower.

The data used in this report have been collected and culled by the Office of Nuclear Waste Isolation's contractor, Stone & Webster Engineering Corporation (SWEC). Their refinement of the original sample data included deleting depressed, local grossly overpressured, and underpressured data so that the resulting data base more accurately represented the original flow system at steady-state conditions before oil and gas production (Bair et al., 1985). A total of 109 data values were available for the Pennsylvanian aquifer, whereas the Wolfcamp retained 85 in the final data set. It is worth noting, at this point, that the sample locations for the two aquifers are quite different.

The first step in any geostatistical analysis is to calculate experimental semivariograms and fit (if possible) a valid model to the graph. In the multivariate case, univariate semivariograms must be constructed for each aquifer and a cross semivariogram between the two sets of potentiometric data.

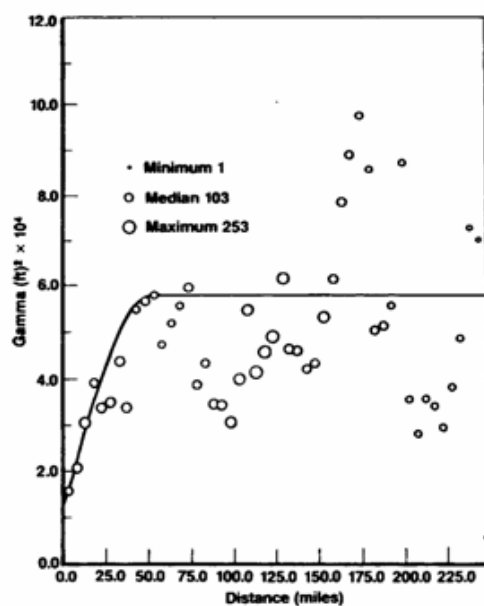


Figure 2. Semivariogram model for Pennsylvanian data.

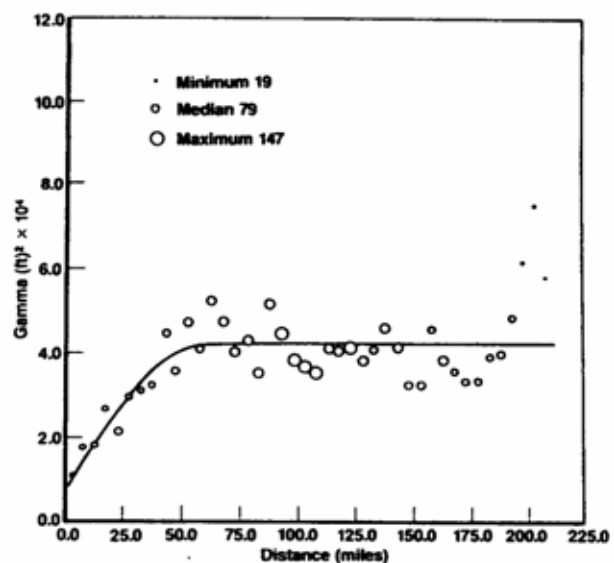


Figure 3. Semivariogram model for Wolfcamp data.

Both aquifers exhibit very strong trends, as would be expected with this type of variable. For both the Pennsylvanian and the Wolfcamp aquifers, a linear trend was sufficient to describe the study area. Experimental semivariograms are calculated on the residuals from these trends. Figures 2 and 3 show the calculated semivariogram graphs for the residual data in the two aquifers and the models fitted to them. Comparing the model with the data in each case, it can be seen that the model tends to be a little higher than one would intuitively expect. This is a function of the trend surface fitting, which tends to result in a downward bias in the observed variation between sample values. The "true" sill is determined by a cross-validation process using universal kriging (UK) on the original data value. This process consists of removing a data

value from the sample set, estimating the value at that location from the surrounding samples, and comparing the estimate with the “true” value. This procedure is repeated for every data value. In ideal circumstances, the resulting statistics should average zero and have a standard deviation of unity.

The models fitted to these graphs are both simple spherical ones. The Pennsylvania residuals show a range of influence of 50 miles, with a nugget effect of 12,000 ft² and a sill (for the spherical component) of 46,000 ft². The Wolfcamp has a slightly longer range at 60 miles, a nugget effect of 11,000 ft², and a sill of 34,000 ft².

Because there are only two variables considered in this study, one cross semivariogram needs to be constructed. Figure 4 shows the resulting experimental semivariogram when taking the difference between the Pennsylvania and the Wolfcamp data. It should be remembered that the data being used in this calculation are the residuals from a linear trend in each case. A model has been fitted to the graph. It is spherical with a range of 35 miles, a nugget effect of 15,000 ft² and a sill of 28,000 ft².

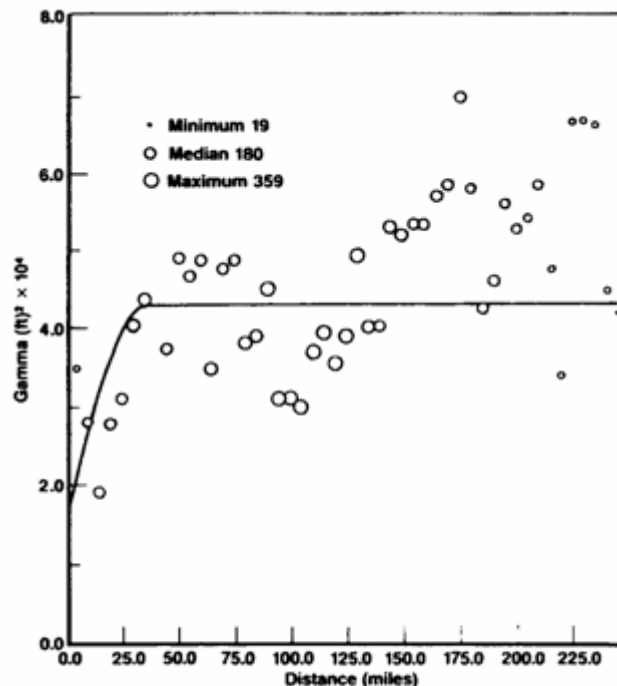


Figure 4. Cross variogram for Pennsylvania and Wolfcamp data.

A cross-validation exercise was carried out using the above models and the multivariate universal co-kriging (MUCK) system. In this study, an average of zero and a standard deviation of 1.13 were obtained. These figures were considered acceptable for the purposes of this study.

The second stage of the analysis is to produce estimates of unsampled locations for each variable. To illustrate this process, a square grid of points has been estimated right across the study area. This grid can be used to produce maps for the potentiometric surface of each aquifer. Because a kriging technique, a standard error is obtained for each grid point, enabling the construction of maps of the “reliability” of the estimation of the potentiometric level.

Figure 5 shows the MUCK-generated map for estimating the potentiometric head values in the Pennsylvanian aquifer. This estimation is carried out using both Pennsylvanian and Wolfcamp samples, and this is emphasized by the posting of the sample locations, showing intersections in the Pennsylvanian as solid circles and those in the Wolfcamp as stars.

For comparison, Figure 6 shows the map estimated by universal kriging using only sample values for the Pennsylvanian aquifer. Visual comparison of the two maps shows local variation in many of the contours, particularly in those areas where the additional sampling is included by the MUCK estimation. This is particularly noticeable in the northeastern Panhandle and the central portion of the study area. The contours change very little in those areas where the Wolfcamp samples merely augment the existing Pennsylvanian sampling.

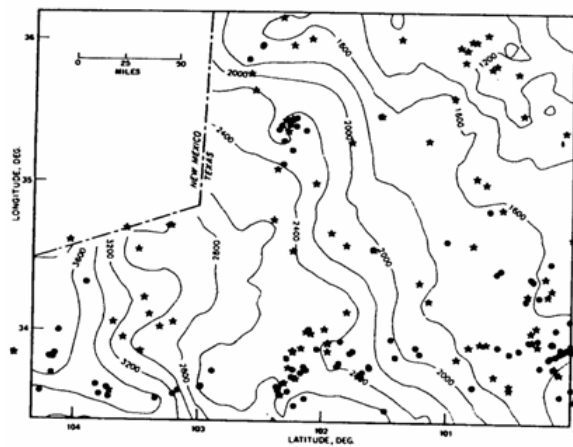


Figure 5. Co-kriging estimates for potentiometric level—Pennsylvanian aquifer.

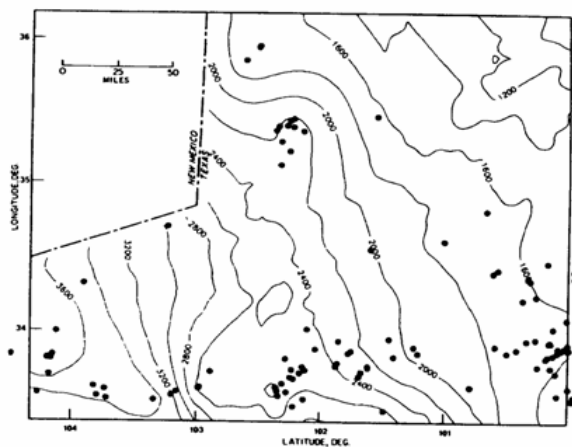


Figure 6. Kriging estimates for potentiometric level—Pennsylvanian aquifer.

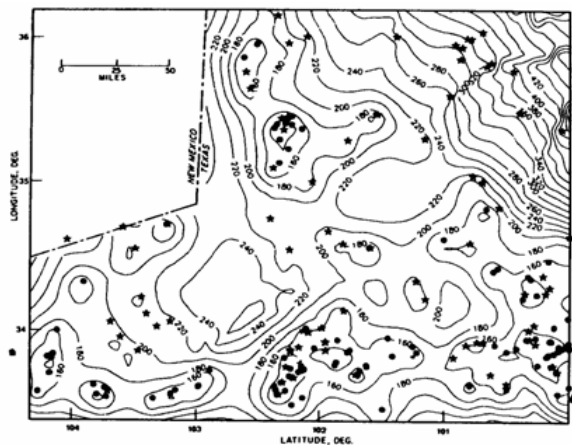


Figure 7. Co-kriging standard errors for potentiometric level—Pennsylvanian aquifer.

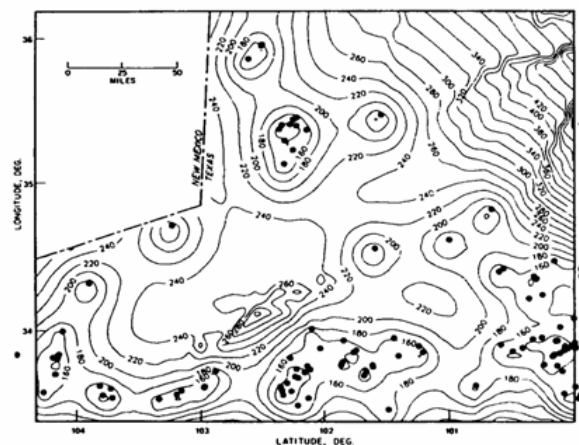


Figure 8. Kriging standard errors for potentiometric level—Pennsylvanian aquifer.

Figures 7 and 8 show the corresponding “standard error” maps for Figures 5 and 6, respectively. Visual comparison of this pair of maps reinforces the interpretation above. The inclusion of the Wolfcamp samples significantly reduces the standard errors in those areas where Pennsylvanian samples are scarce.

The MUCK system simultaneously produces estimates standard errors for the Wolfcamp aquifer while mapping the Pennsylvanian. Figures 9 and 11 show the estimated and standard error maps respectively. A separate universal kriging exercise must be undertaken to produce Figures 10 and 12 for the Wolfcamp alone.

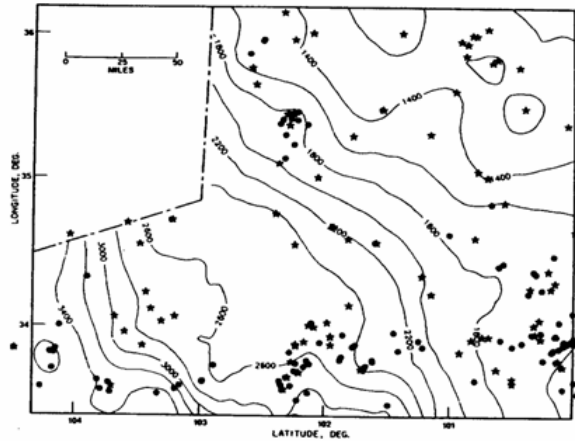


Figure 9. Co-kriging estimates for potentiometric level—Wolfcamp aquifer.

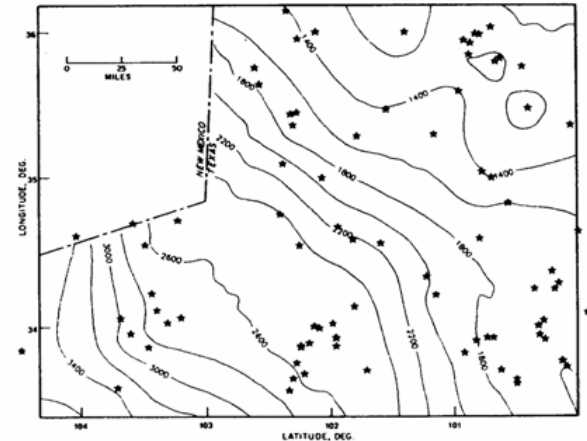


Figure 10. Kriging estimates for potentiometric level—Wolfcamp aquifer.

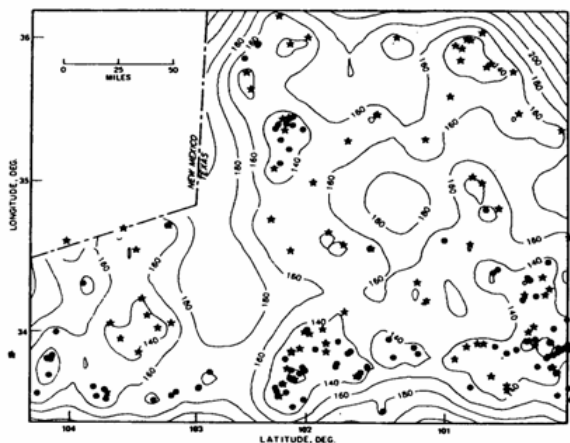


Figure 11. Co-kriging standard errors for potentiometric level—Wolfcamp aquifer.

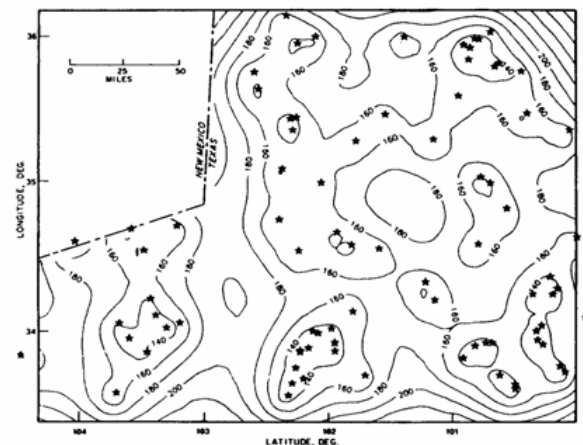


Figure 12. Kriging standard errors for potentiometric level—Wolfcamp aquifer.

Visual inspection of each of these pairs of maps reveals that the incorporation of additional sampling, albeit for a different variable, can change the estimation of the potentiometric surface. However, this does not provide a quantitative measure of what, if any, advantage has been gained by using MUCK rather than the more conventional UK. In theory, MUCK should provide smaller standard errors than UK because more information is used in the production of the estimates, both in terms of numbers of data and in the additional modelling required for the cross semivariogram. If the multivariate approach has significantly improved the errors, the result will be a tighter confidence banding around the estimates of the surfaces.

The two methods were compared as follows. The standard error at each of the grid points was squared to provide the estimation variance. For each grid point, then, there is a univariate universal kriging estimation variance (UKEV) found by UK and a multivariate estimation variance found by MUCK. The UKEV is divided by the MUCK estimation variance to produce a variance ratio between the two methods. A variance ratio of exactly 1 indicates a location where the multivariate approach adds no extra reliability. Any value greater than 1 indicates a location where MUCK has improved the estimate.

Figure 13 shows a contour map of the variance ratio between UK and MUCK for the Pennsylvanian data. Both Pennsylvanian and Wolfcamp sample locations are shown on the map (circles and stars, respectively) to emphasize the multivariate nature of the MUCK procedure. It can be seen that the highest ratios are in locations where the Wolfcamp samples fill gaps left in the Pennsylvanian sampling. This map illustrates the reduction in estimation variance achieved by including sample data obtained from a different, but significantly correlated, variable. The variance ratios range from 1.0 to over 1.8, with an average over the study area of 1.15. A simplistic interpretation of this figure would be that the average increase in “efficiency” of the estimates is around 15%. Perhaps more accurately, we should take the square root of this figure and expect a general decrease in “standard error” of around 7.5%.

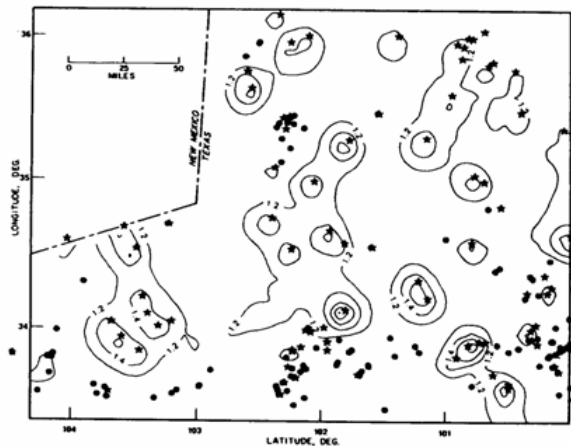


Figure 13. Ratio between kriging and co-kriging variances—Pennsylvanian aquifer.

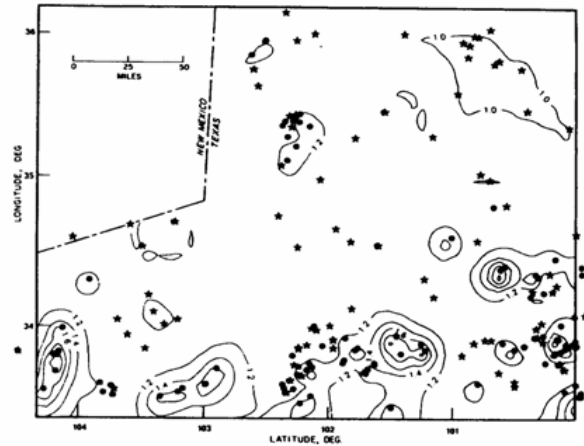


Figure 14. Ratio between kriging and co-kriging variances—Wolfcamp aquifer.

Figure 14 shows the corresponding variance ratio map for the Wolfcamp aquifer. Again, it can be seen that “improvement” is most noticeable at those locations where Pennsylvanian samples exist in gaps in the sampling pattern. Here, the average variance ratio is 1.09, resulting in approximately 4% general improvement in the final confidence levels.

This case study shows that the estimation of a spatial variable can be enhanced using samples from another correlated variable sampled at different locations. It should be remembered that any reduction in uncertainty at this stage of estimation of the hydrogeologic variables will carry through to any later sensitivity and/or simulation analyses.

Summary

It has been suggested that estimation of spatially distributed variables might be improved if cross correlations between variables could be included in the estimation process. Traditional co-kriging has been considered and found to be inapplicable to the problem under study. A new formulation of the cross semivariogram is proposed that not only is applicable to this particular study, but also avoids many of the problems associated with previous co-kriging methods.

The new cross semivariogram model is used to produce a minimum variance unbiased estimation technique that has been dubbed multivariate universal co-kriging, partly to distinguish it from traditional co-kriging. The case study for the Texas Panhandle illustrates that MUCK will improve estimation to some extent throughout the study area but will have significant impact on those areas that have been undersampled for a particular variable.

Literature Cited

1. Bair, E. S., O'Donnell, T. P., and Picking, L. W., 1985. *Hydrogeologic Investigations Based on Drill-Stem Test Data: Palo Duro Basin, Texas and New Mexico* BMI/ONWI-566, February, prepared by Stone & Webster Engineering Corporation for Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, Ohio.
2. Carr, J. R., and McCallister, P. G., 1985. "An Application of Co-Kriging for Estimation of Tripartite Earthquake Response Spectra", *Mathematical Geology*, 17, 5, 527-545.
3. Carr, J. R., Myers, D. E., and Glass, C. E., 1985. "Co-Kriging - A Computer Program", *Computers and Geosciences*, 11, 2, 111-127.
4. Harper, W. V., and Furr, J. M., 1986. Geostatistical Analysis of Potentiometric Data in the Wolfcamp Aquifer of the Palo Duro Basin, Texas, BMI/ONWI-587, April, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, Ohio.
5. Harper, W. V., Basinger, K. V., and Furr, J. M., 1986. "Geostatistical Analysis of Potentiometric Data in the Pennsylvanian Aquifer of the Palo Duro Basin, Texas, BMI/ONWI-680, January, 1988, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, Ohio.
6. Guarascio, M., 1976. "Improving the Uranium Deposit Estimations (the Novazza Case)", in *Advanced Geostatistics in the Mining Industry*, M. Guarascio et al. (Eds.), D. Reidel, 351-367.
7. Matheron, G., 1971. *The Theory of Regionalized Variables and Its Applications*, Cahier No. 5, Centre de Morphologie Mathematique de Fontainebleau, 211 pp.
8. Myers, D. E., 1982. "Matrix Formulation of Co-Kriging", *Mathematical Geology*, 14, 3, 249-257.
9. Myers, D. E., 1983. "Estimation of Linear Combinations and Co-Kriging", *Mathematical Geology*, 15, 5, 633-637.
10. Myers, D. E., 1984. "Co-Kriging-New Developments", in *Geostatistics for Natural Resources Characterization*, ed. G. Verly et al. (Eds.), D. Reidel, 295-305.
11. Unal, A., and Haycocks, C., 1986. "Geostatistical Coal Resource Characterization Using Correlations", in *Proc. 19th APCOM Symposium*, Pennsylvania State University, April, SME-AIME (publ.), 231-239.

Mathematical Appendix

Ordinary kriging (OK)

To estimate the value, T , at an unsampled location A , we form a linear estimator:

$$T^* = \sum w_i g_i$$

where g_i is the observed value for sample S_i , w_i is the weighting factor given to sample i , and T^* denotes an estimate of T .

The error incurred in the estimation may be expressed as $(T - T^*)$. It can be shown that (in the absence of trend):

$$E[T - T^*] = 0 \text{ if } \sum w_i = 1$$

The variance of the estimation errors then becomes

$$\sigma_e^2 = E[(T - T^*)^2]$$

which, with a little algebraic manipulation, reduces to

$$\sigma_e^2 = 2 \sum w_i \gamma(S_i, A) - \sum \sum w_i w_j \gamma(S_i, S_j) - \gamma(A, A)$$

where $\gamma(A, S_i)$ is the semivariogram value between the unsampled location and the i th sample, and so on.

Minimization of this variance with respect to the weighting factor, w_i , and subject to the unbiasedness constraint results in a system of equations as follows:

$$\sum w_j \gamma(S_i, S_j) + \lambda = \gamma(S_i, A) \text{ for } i = 1, 2, 3 \dots n$$

$$\sum w_j = 1$$

This is the ordinary kriging system. The use of these weighting factors also results in a simplification of the estimation variance:

$$\sigma_{ok}^2 = \sum w_i \gamma(S_i, A) + \delta - \gamma(A, A)$$

Universal kriging (UK)

The universal kriging system is similar in form, with extra Lagrangian multipliers (λ) added for each term in the trend. For example, if a linear trend existed in a two-dimensional plane this would be expressed as

$$m(x, y) = b_0 + b_1 x + b_2 y$$

Knowledge of the b coefficients is not necessary for the kriging stage, only the form of the trend. The constraints for unbiasedness then become

$$\sum w_j = 1$$

$$\sum w_j x_j = x_A$$

$$\sum w_j y_j = y_A$$

and $\sum w_j \gamma(S_i, S_j) + \lambda_0 + \lambda_1 x_i + \lambda_2 y_i = \gamma(S_i, A) \text{ for } i=1, 2, 3 \dots n$

The universal kriging variance simplifies to

$$\sigma_{uk}^2 = \sum w_i \gamma(S_i, A) + \lambda_0 + \lambda_1 x_A + \lambda_2 y_A - \gamma(A, A)$$

Multivariate universal co-kriging (MUCK)

We illustrate this approach with a three-variable example. Suppose we wish to estimate the value of variable g at location A . We have n observations on g available in the immediate neighborhood. We also have m observations on another variable, f , and t observations on a third variable, r . We extend the above notation so that (for example) S_{fi} denotes the i th sample with a measured value of variable f and A_g the value of variable g at location A .

Our estimator now becomes

$$T^* = \sum w_i g_i + \sum v_j f_j + \sum u_k r_k$$

One set of constraints resulting in an unbiased estimator (in the absence of trend) is

$$\sum w_i = 1 \quad \sum v_j = 0 \quad \sum u_k = 0$$

Of course, if we were estimating the value of variable f at this location, the $\{w_i\}$ would sum to zero and the $\{v_j\}$ would sum to 1, etc.

The general form of the estimation variance becomes

$$\begin{aligned} \sigma_e^2 = & 2 \sum w_i \gamma(S_{gi}, A_g) + 2 \sum v_j \gamma(S_{fj}, A_g) + 2 \sum u_k \gamma(S_{rk}, A_g) \\ & - \sum \sum w_i w_p \gamma(S_{gi}, S_{gp}) - 2 \sum \sum w_i v_q \gamma(S_{gi}, S_{fq}) \\ & - \sum \sum v_j v_q \gamma(S_{fj}, S_{fq}) - 2 \sum \sum w_i u_t \gamma(S_{gi}, S_{rt}) \\ & - \sum \sum u_k u_t \gamma(S_{rk}, S_{rt}) - 2 \sum \sum v_j u_t \gamma(S_{fj}, S_{rt}) - \gamma(A_g, A_g) \end{aligned}$$

Minimization of this variance subject to the unbiasedness constraints gives

$$\begin{aligned} & \sum w_i \gamma(S_{gi}, S_{gp}) + \sum v_j \gamma(S_{fj}, S_{gp}) + \sum u_k \gamma(S_{rk}, S_{gp}) \\ & \quad + \lambda_g = \gamma(S_{gp}, A_g) \\ & \sum w_i \gamma(S_{gi}, S_{fq}) + \sum v_j \gamma(S_{fj}, S_{fq}) + \sum u_k \gamma(S_{rk}, S_{fq}) \\ & \quad + \lambda_f = \gamma(S_{fq}, A_g) \\ & \sum w_i \gamma(S_{gi}, S_{rt}) + \sum v_j \gamma(S_{fj}, S_{rt}) + \sum u_k \gamma(S_{rk}, S_{rt}) \\ & \quad + \lambda_r = \gamma(S_{rt}, A_g) \\ & \sum w_i = 1 \\ & \sum v_j = 0 \\ & \sum u_k = 0 \end{aligned}$$

and a simplified co-kriging variance of

$$\sigma_{ock}^2 = \sum w_i \gamma(S_{gi}, A_g) + \sum v_j \gamma(S_{fj}, A_g) + \sum u_k \gamma(S_{rk}, A_g) + \lambda_g - \gamma(A_g, A_g)$$

With this formulation it is fairly simple to generalize to the universal co-kriging form, including the trend (if any) for each variable as extra constraints. We find that this formulation also highlights terms that remain constant, regardless of which variable is being estimated, and enables us to proceed to the multivariate case with ease.

For matrix formulation of these equations, see Myers (1982).