GEOSTATISTICAL AND

STOCHASTIC INVERSE ANALYSIS

OF THE WOLFCAMP AQUIFER

William V. Harper
Battelle Project Management Division
Columbus, Ohio 43201

Joseph L. Devary
Battelle Project Management Division
Richland, Washington 99352

ABSTRACT

The quantification of uncertainty is critical in both the performance assessment and licensing of a high level waste repository. Under expected conditions, repository performance is dominated by continuous processes with far field flow being one of these key processes. We present a kriging analysis of potentiometric data in the Wolfcamp aquifer of the Palo Duro Basin in northern Texas combined with a stochastic inverse analysis using adjoint sensitivity theory. These results may be used to identify additional data needs. A careful development of the steps of the kriging analysis is presented. Universal kriging which combines fitting a polynomial trend surface with ordinary kriging is shown to provide a good fit to the isotropic potentiometric data. The necessary adjoint sensitivity theory is applied to backsolve for the hydraulic conductivities via an iterative stochastic inverse method.

INTRODUCTION

The Office of Nuclear Waste Isolation (ONWI) has ongoing performance assessment analyses for several potential sites for a high level nuclear waste repository in salt. An understanding of the far field hydrologic system is a key part of such studies. While deterministic analyses play an important role, it is critical that detailed sensitivity/uncertainty analyses be performed to identify additional data needs and to demonstrate compliance with federal regulations. This report combines geostatistical analysis, adjoint sensitivity theory, and nonlinear optimization techniques to solve what is called the stochastic inverse problem.

In the usual or forward solution of the far field flow system of equations, hydraulic conductivities and boundaries conditions are input and the potentiometric head surface is the resulting solution of these equations. In most cases there is better available potentiometric head data than available hydraulic conductivity data. The purpose of the stochastic inverse technique is to take advantage of this and use the available head data to backsolve for the conductivity values that result in the best fit. This is really a calibration or tuning of the conceptual model. Rather than letting the modeler arbitrarily adjust the parameters, the stochastic inverse solution automatically best fits the conductivities in an iterative production combining geostatistics and adjoint sensitivity theory.

Geostatistical Analysis of the Wolfcamp Aquifer

The Wolfcamp aquifer underlies a potential high level nuclear waste repository in the Permian Basin. Deaf Smith county in northern Texas is one of the sites being studied for such a nuclear waste facility. The Wolfcamp aquifer plays an important part in any study of far field flow in this area. Groundwater travel times and travel paths and their associated uncertainty are key performance measures in an overall performance assessment uncertainty analysis. As part of this uncertainty analysis, the potentiometric head surface of the Wolfcamp must be quantified. The geostatistical analysis given in this report is one method of estimating both that surface and its uncertainty. Figure 1 shows the portion of Texas that this study focuses on.

The location and values of the 85 potentiometric data values are seen in Fig. 2. A rough examination of these values indicates a general trend of high to low values in a northeastern direction. The first step in the analysis involves the creation of directional semi-variograms. These will be used to check for anisotropy and trend. Figure 3 shows the semi-variograms for the four major directions. Each point on this plot represents all the pairwise combinations in the given direction that fall within a

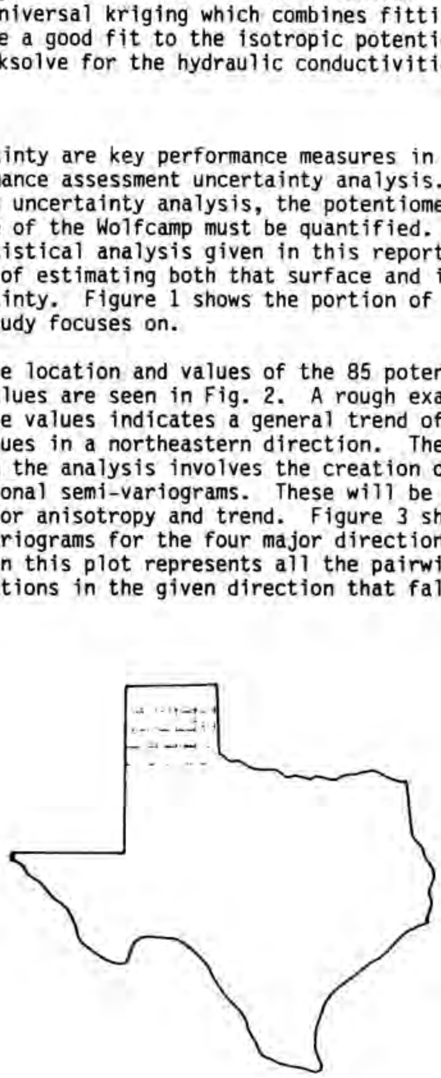

Fig. 1. Location Map For The Study Region.

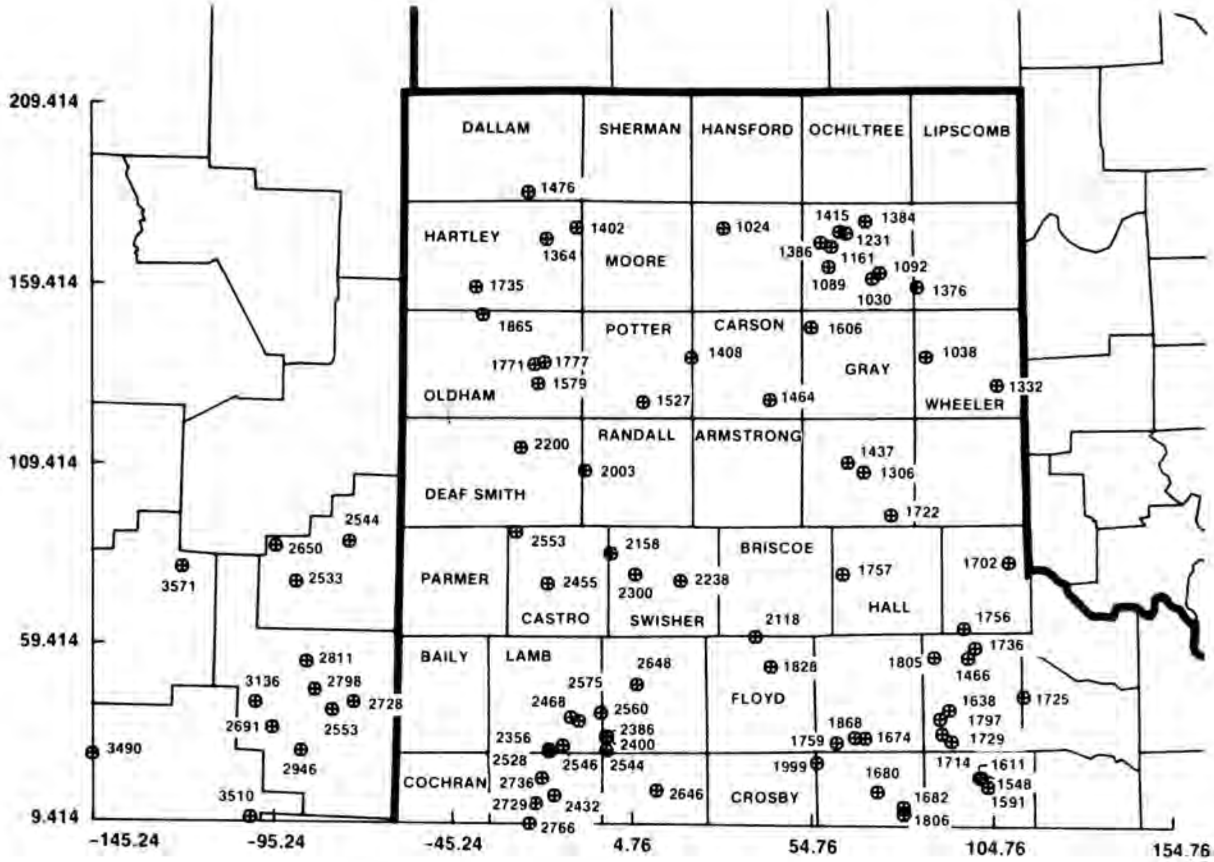

Fig. 2. Plot of Potentiometric Database.

particular five mile cell. For example the square just beyond the 100 mile label on the horizontal axis represents half the average squared difference between all pairs of points in the northwest-southeast direction that are from 100 to 105 miles apart. They are plotted at the average distance between the points (somewhere between 100 and 105 for this particular example). There is a wealth of information in Fig. 3. The four semi-variograms appear to be almost identical for about 75 miles, that is they exhibit isotropy up to this point. After 75 miles a spreading of the plots occurs. This could be the result of either an anisotropy or trend; however, this appears to be a fairly clear case of trend coming into play (this will be investigated below). If it is indeed the result of a trend dominating the semi-variograms after 75 miles, then the direction of this trend is primarily in the southwest-northeast and east-west directions (as seen by the increasing semi-variograms) with the contours running primarily in the northwest-southeast direction (as indicated by the relatively horizontal pattern of the semi-variogram). If the kriging neighborhood used is to be less than 75 miles the assumption of isotropy might be reasonable; however, a better model may result from taking the trend into account. Figure 4 provides an even more graphic depiction of the trend. This is a shading plot and clearly shows the general pattern not seen quite as clearly in Fig. 3. Since geostatistical analysis is based on visual pattern recognition, shading plots such as this are very useful. Figure 5 is an overall semi-variogram plot with all directions combined. It illustrates the average behaviour seen in the directional semi-variograms of Fig. 3. The evidence of trend can be seen in this overall plot also.

Statistical linear regression techniques were used to fit a trend surface to the 85 potentiometric values using the statistical software package SAS (SAS, 1982). This analysis shows that a linear trend provides a good fit to the data while a quadratic surface does not add anything significant. Thus we may infer that the drift or trend of the data is

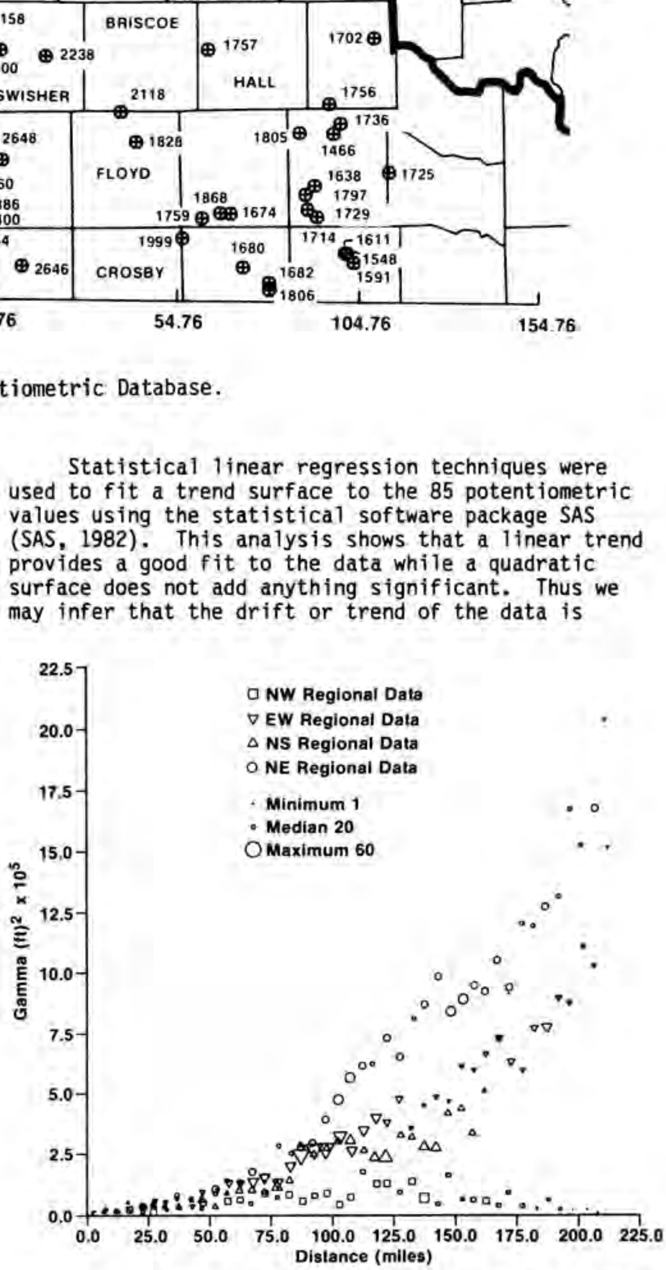

Fig. 3. Directional Semi-Variograms on Original Data.

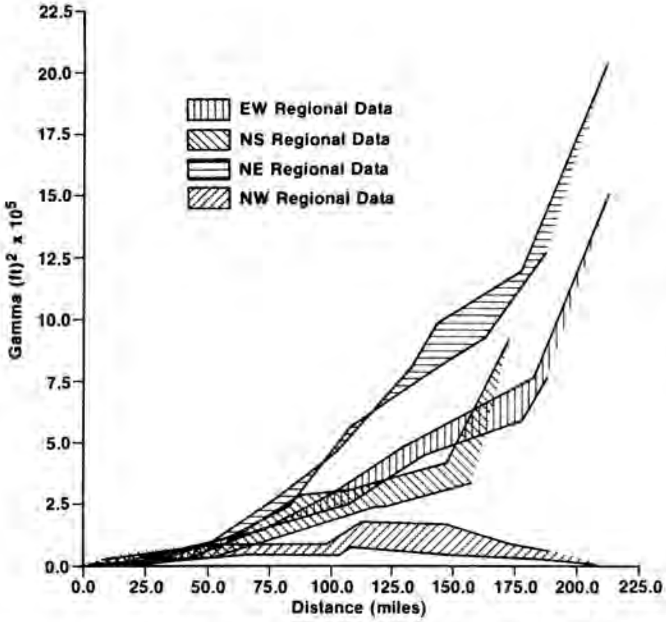

Fig. 4. Shading Plots For Directional Semi-Variograms in Fig. 3.

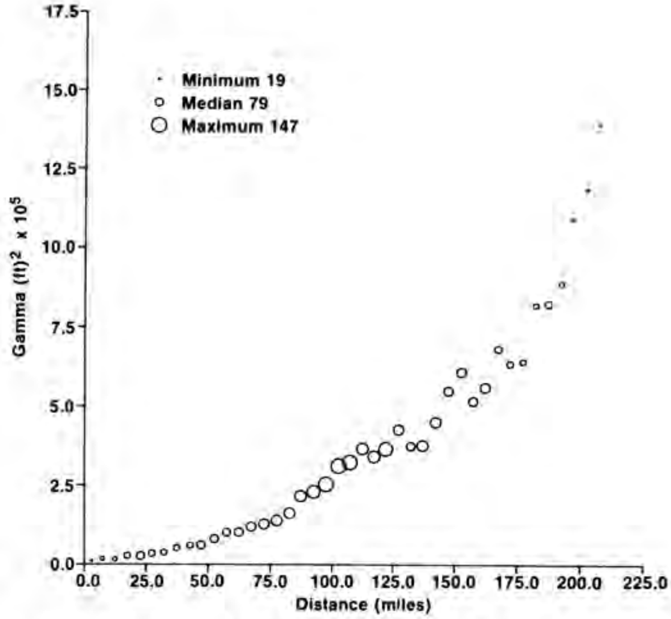

Fig. 5. Overall Semi-Variogram On Original Data.

linear. The residuals (difference between the predicted and observed potentiometric values) of the linear trend surface are then subjected to semi-variogram analysis as were the original data in the previous paragraph. Figure 6 gives the directional semi-variograms for the residuals of the linear trend surface. This figure no longer shows an obvious trend and the overall semi-variogram pattern is the same in all four directions. The shading plot of Fig. 7 shows this clearly. At this point we can safely assume that the data is isotropic once the linear trend is accounted for. This implies the same spatial variability in all directions and allows us to model one semi-variogram for all directions.

The overall semi-variogram for the linear trend surface residuals is seen in Fig. 8. Plotted on top of the empirical semi-variogram data is a thereotical model. It is the theoretical model that will be used in the actual kriging interpolation. Finding a good model for the semi-variogram involves an iterative procedure in which various theoretical models are tried and refined until the best visual fit is obtained between the model and empirical semi-variogram. Some software packages allow a black box approach where one never has to visually match the selected model to the actual data. Such approaches should be avoided as they can lead to poor results. While statistical methods are important in picking a theoretical semi-variogram model, they are secondary to a good visual fit. Unfortunately, the statistical methods used to identify models that represent the spatial variability (be they semi-variograms or generalized covariances) are not totally reliable. The particular model portrayed in Fig.8 is a spherical semi-variogram with a nugget of 7,000 ft^2, sill of 35,000 ft^2, and a range of 60 miles. The nugget of 7,000 corresponds to a measurement error of 84 feet (square root of 7,000) in the potentiometric data which is in agreement with field experience that estimates head measurement error to be roughly 100 feet.

Once a theoretical semi-variogram model is fit, statistical cross-validation may be performed. This will help check the adequacy of the proposed model and also identify those data values that do not fit the

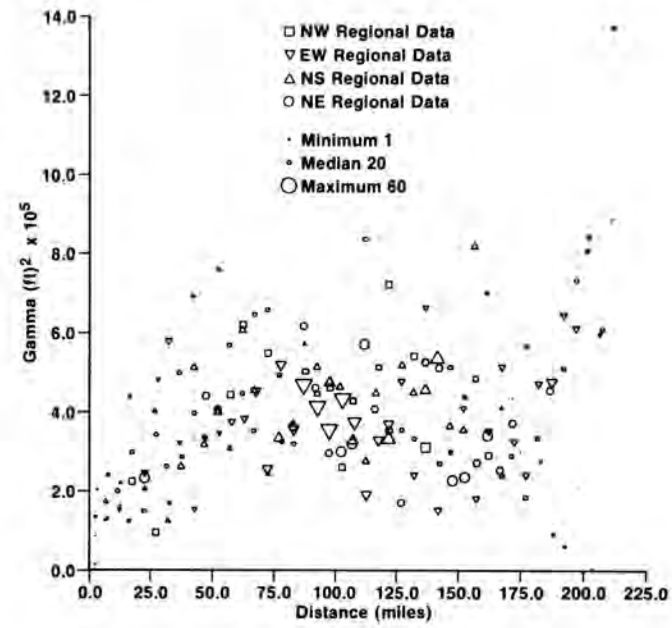

Fig. 6. Directional Semi-Variograms On Linear Residuals.

overall pattern of the rest of the data. Data identified in this manner require close scrutiny, but are not deleted from the data set for purely statistical reasons. There must be sound traceable geologic evidence before the questionable data is removed from the data set. Cross-validation removes each data value one at a time from the data set, and estimates its value from the surrounding neighbors using kriging with the proposed theoretical semi-variogram model. The cross-validation residuals (difference between the predicted and actual value) are then normalized by their kriging standard error and should have certain statistical properties if the model provides a good fit. The standardized residuals should have a mean of 0 and a standard deviation of 1.

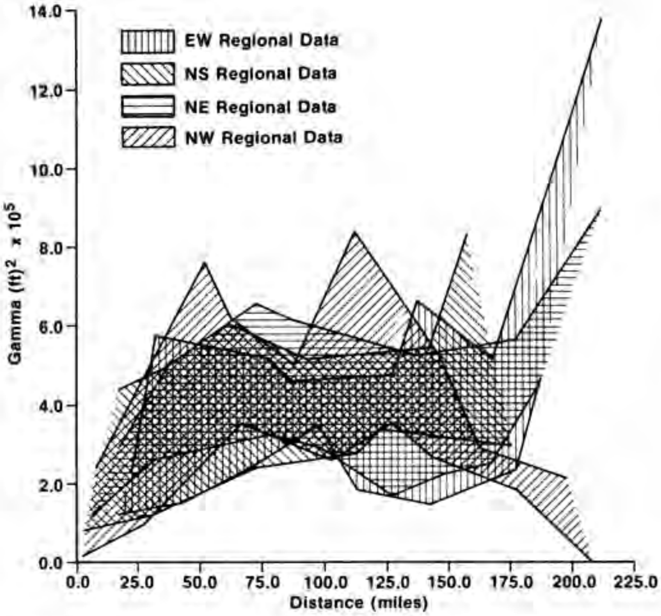

Fig. 7. Shading Plots for Directional Semi-Variograms in Fig. 6.

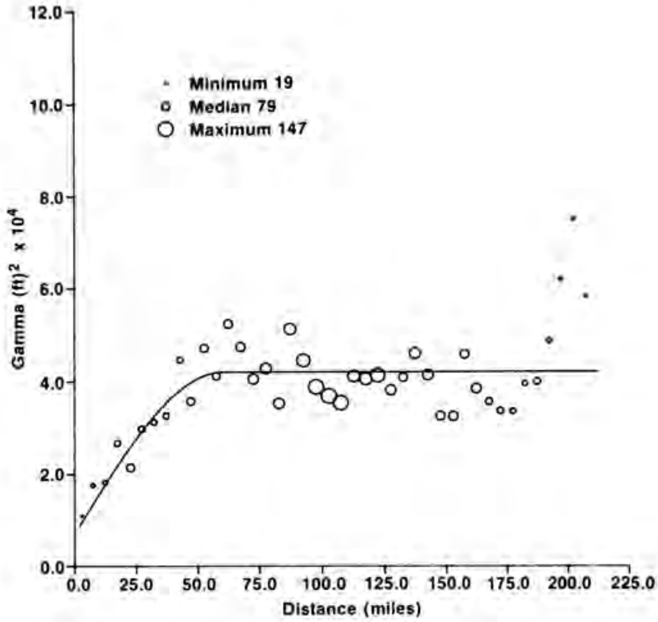

Fig. 8. Overall Semi-Variogram With Superimposed Theoretical Model.

Minor deviations from this are not a cause for concern. Cross-validation statistics on the Wolfcamp potentiometric data for the semi-variogram in the prior paragraph resulted in a mean of .01 and a standard deviation of 1.17 which are reasonable values. The value of 1.17 indicates that there are probably a few residuals larger than expected but not enough to cause a poor fit. Regardless of this one should always plot the standardized residuals (cross-validation residuals divided by their kriging standard error) and study them. Identification of potentially anomalous data is thus possible. Plotting also will allow one to identify why a large standardized residual occurred (the observed value will in some way vary considerably from its nearest neighbors). Figure 9 is the plot of the standardized residuals for the above theoretical semi-variogram model. There are no residuals larger than 3 which is good; however, there are two residuals greater than 2.5. One of these is in New Mexico and would have minimal impact on the potentiometric surface. Its value of 2,691 is lower than would be expected given its neighbors. The second one is in Gray county and has a value of 1,606 which is larger than would be expected based on its neighbors and the assumed theoretical semi-variogram model. This value does have some impact on the possible travel path from a location in Deaf Smith county toward the northeast (perpendicular to the contours seen in Fig. 10), but is not a major deviation from the overall potentiometric surface.

After the semi-variogram model has been selected, it is a straightforward process to produce a kriged surface for both the expected potentiometric surface and its associated uncertainty surface. In Fig. 10 the trend is easily seen to run primarily from the west to the northeast or east direction. This is in the direction of the Amarillo uplift. The expected surface is similar to others produced in other reports (e.g., see BMI/ONWI-515). While Fig. 10 provides important information, perhaps the most important results may be seen in Fig. 11. This figure shows the kriging standard error contours. These may be used to form confidence intervals for the contour surface given on the previous page. Thus ONWI can not only predict the potentiometric surface but may also quantify the reliability in those predictions.

Stochastic Inverse Using Adjoint Sensitivity Theory

A two-dimensional version of the finite element code CFEST was used to model the far field flow system for the Wolfcamp aquifer. The code was adjointed for steady state flow. This adjointed version of CFEST allows a thorough sensitivity analysis of selected performance measures. For the stochastic inverse solution, the desired performance measure is a weighted sum of the squared deviations of predicted minus observed potentiometric head values.

The mathematics for the solution of the stochastic inverse problem is based on the results of S. Neuman (1982). Initially nine hydraulic conductivity zones were established with initial (or prior) values. The recharge values necessary to set the proper potentiometric head values were found via the iterative solution that resulted in updated or posterior conductivity values. Given the initial values of hydraulic conductivity and observed head data, an iterative procedure using the kriging covariance matrix, adjoint sensitivity, and nonlinear parameter estimation techniques was used to update the hydraulic conductivity and recharge values in each iteration. These updated values become the prior values in the next iteration. This process continued until convergence was obtained. At this point, a finer 18 region zonation scheme was used taking advantage of the results obtained for the 9 zone system.

Although the inferred hydraulic conductivity values appear to be consistent with the known geologic structure of the Palo Duro Basin, further evaluation of this analysis and existing permeability data is necessary before a great deal of confidence may be given to the conceptual model. This stochastic inverse analysis is still considered preliminary. The uncertainty in the recharge, hydraulic conductivity, and held head conditions will be propagated to both travel path and travel time uncertainties in the future.

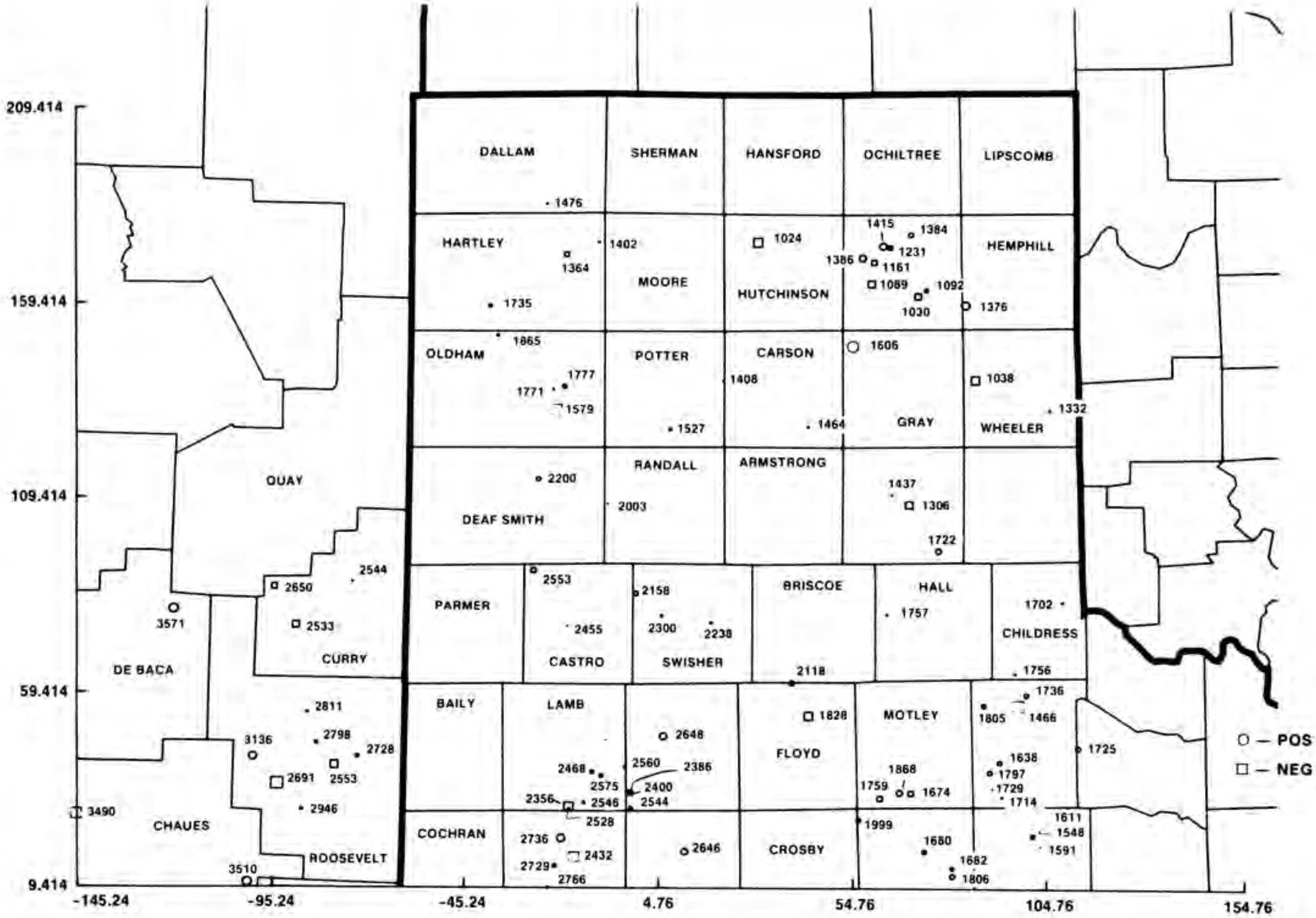

Fig. 9. Standardized Residual Plot.

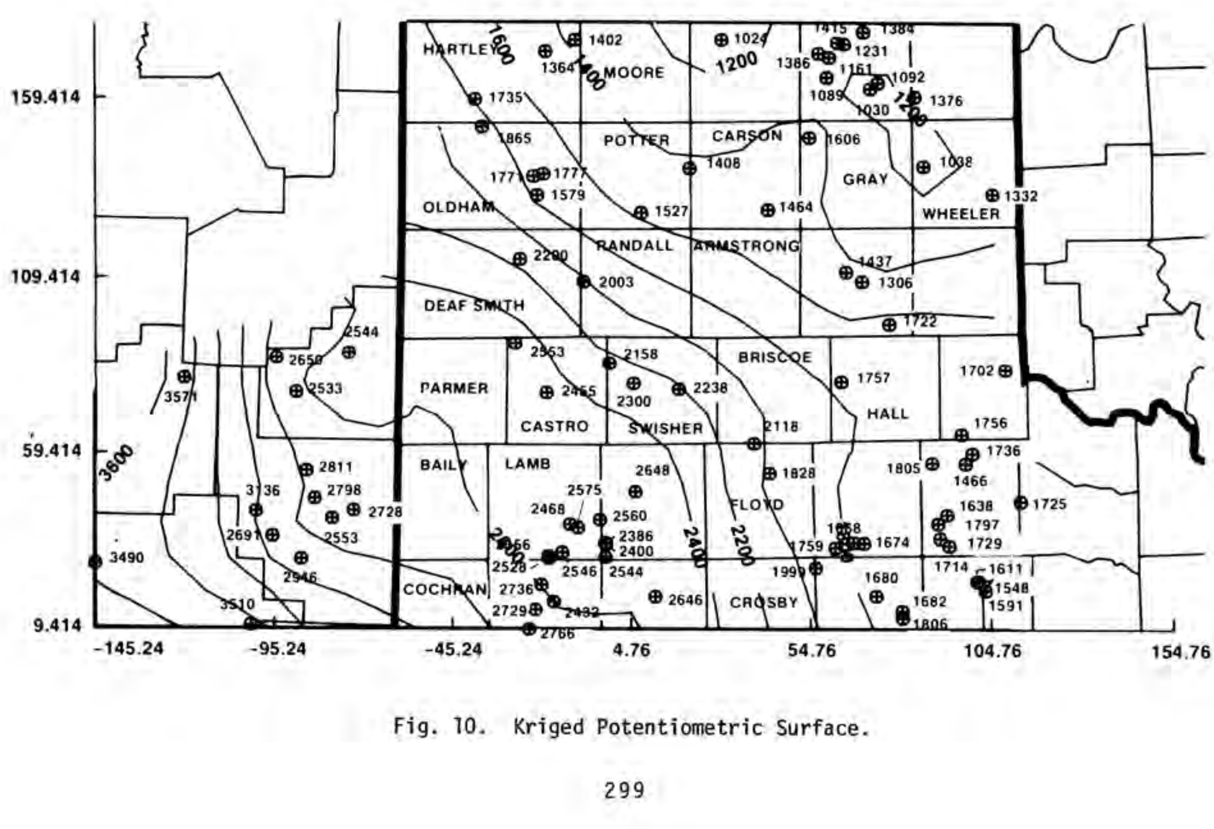

Fig. 10. Kriged Potentiometric Surface.

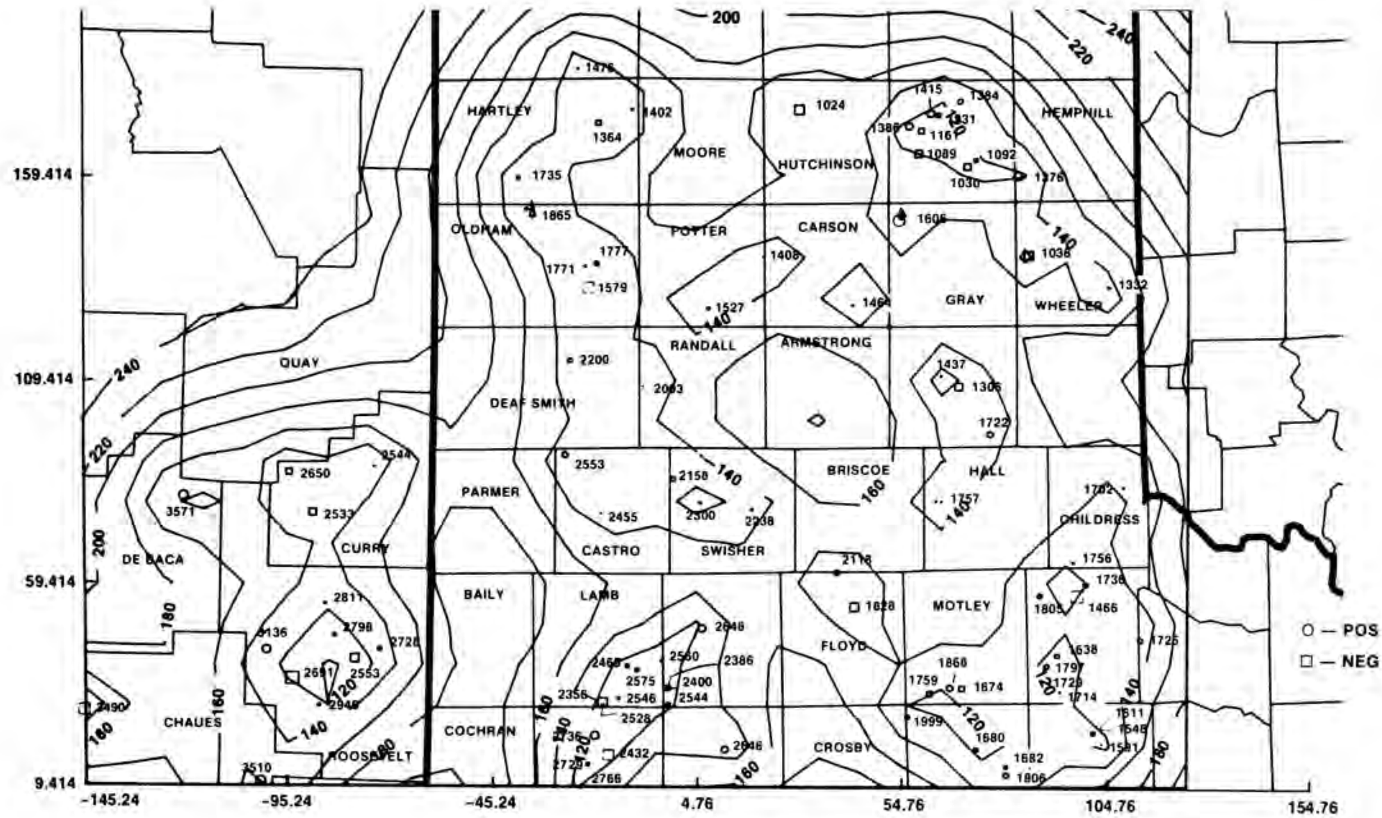

Fig. 11. Kriged Standard Error Surface.

Conclusion

The stochastic inverse procedure used in the paper was valuable in calibrating the 2-D far-field flow model illustrated here. It provides an automatic fine tuning of the conceptual model that will best represent the available potentiometric data. This is an important phase in the overall sensitivity/uncertainty analyses of a far field flow system, and will lead to estimates of uncertainty in groundwater travel times and travel paths.

References

J. L. DEVARY, W. V. HARPER, J. F. SYKES, and J.L. WILSON "Far Field Flow Uncertainty Analysis for the Palo Duro Basin," Scientific Basis for Nuclear Waste Management VII, G. L. McVay, Editor, Materials Research Society Symposia Proceedings, Vol. 26, pp. 397-404, Elsevier Science Publishing Co., Inc., New York (1984).

INTERA, Adjoint Sensitivity Theory for Steady-State Ground-Water Flow, BMI/ONWI-515, November, 1983.

S. P. NEUMAN, "Statistical Characterization of Aquifer Heterogeneities: An Overview," Geologic Society of America, Special Paper 189 (1982).

J. L. WILSON, and W. V. HARPER, "Evaluation of Prediction Uncertainty in Performance Assessment: Far Field Flow at Bedded Salt Sites," Proceedings of the 1983 Civilian Radioactive Waste Management Information Meeting, CONF-831217, pp. 337-346, United States Department of Energy, Office of Civilian Radioactive Waste Management, Washington, D.C. (February 1984).