CONF-910213--10

SAND--90-1056C

DE91 000405

OCT 0 5 1990

PARAMETER ESTIMATION TECHNIQUES AND UNCERTAINTY IN GROUND WATER FLOW MODEL PREDICTIONS†

ZIMMERMAN, D. A* AND P. A. DAVIS**
*GRAM, Incorporated, Albuquerque, New Mexico
**Sandia National Laboratories, Albuquerque, New Mexico

INTRODUCTION

Quantification of uncertainty in predictions of nuclear waste repository performance is a requirement of Nuclear Regulatory Commission regulations governing the licensing of proposed geologic repositories for high-level radioactive waste disposal. One of the major uncertainties in these predictions is in estimating the ground-water travel time of radionuclides migrating from the repository to the accessible environment. The cause of much of this uncertainty has been attributed to a lack of knowledge about the hydrogeologic properties that control the movement of radionuclides through the aquifers. A major reason for this lack of knowledge is the paucity of data that is typically available for characterizing complex ground-water flow systems. Because of this, considerable effort has been put into developing parameter estimation techniques that infer property values in regions where no measurements exist. Currently, no single technique has been shown to be superior or even consistently conservative with respect to predictions of ground-water travel time. This work was undertaken to compare a number of parameter estimation techniques and to evaluate how differences in the parameter estimates and the estimation errors are reflected in the behavior of the flow model predictions. That is, we wished to determine to what degree uncertainties in flow model predictions may be affected simply by the choice of parameter estimation technique used.

OVERALL DESCRIPTION OF THE MODELING PERFORMED

The site used for this study is the Avra Valley aquifer in southeastern Arizona. The basic model is a one-layer, two-dimensional steady-state system; both Dirichlet and Neuman boundary conditions were specified in such a manner that the total flux passing through the system was constant, regardless of the transmissivity distribution used (there were two locations with specified influx and one location with prescribed heads). The only model coefficients that are considered to be uncertain are the transmissivity values; the porosity distribution (a constant value) and all head- and flux-dependent boundary conditions are presumed to be known with certainty. The parameter estimation techniques used included a manual calibration trial-and-error inverse technique, kriging with Generalized Covariances, cokriging, and ordinary kriging combined with linear regression. For each technique applied, multiple realizations of the transmissivity field were generated using the Latin Hypercube sampling technique. A ground-water flow simulation was then performed for each realization using the U.S. Geological Survey's MODFLOW computer code. Estimates of the ground-water travel-time of a conservative tracer moving across the system were made using a particle-tracking technique. The travel time estimates were compiled into a complementary cumulative distribution function (CCDF) for each technique and the CCDF's were compared.

HOW THIS WORK DIFFERS FROM PREVIOUS STUDIES

Previous comparison studies of ground-water flow modeling analyses have typically involved a comparison of results obtained using different uncertainty propagation techniques (e.g., differential analysis versus response surface versus Monte Carlo) usually with a priori knowledge of the input uncertainties. In contrast, this study starts with the raw data and compares results obtained from using different techniques to estimate the model input parameters and their uncertainties while using only one technique, Monte Carlo simulation, to propagate those uncertainties through the model.

† This work was performed at Sandia National Laboratories which is operated for the U.S. Department of Energy under Contract No. DE-AC04-76DP00789

Secondly, unlike most previous studies involving Monte Carlo simulation, in which the results from every run are used in the analysis, we selectively excluded certain runs from the analysis. The choice as to which simulations to include or exclude was based on how well the model predicted heads matched observed values. Simulations in which this match was very poor were considered to have resulted from the use of unrealistic transmissivity realizations which are inconsistent with the chosen conceptual model. To assess the degree to which these unrealistic realizations can affect the uncertainty in model predictions, we computed, for one technique, the CCDF of ground-water travel time using the complete ensemble of realizations and compared these results with the others.

SOME DETAILS OF THE MODELING

A schematic of the flow domain and the boundary conditions used is shown in Figure 1. No areal recharge, ground-water pumpage or streamflow infiltration was used. The gridded domain for the flow modeling was composed of 437 one-mile square finite-difference cells. Some of these cells were combined to form zones of varying size and shape; a single transmissivity value was assigned to all cells within any one zone. The zonation pattern was determined primarily as a function of a soil property classification regarding the fraction of silt and clay particles in the soil. By combining the finite-difference cells into zones, the number of uncertain parameters was reduced from 437 to 215. The kriging methods all produced estimates at the centers of the 437 finite-difference blocks; in multiple-cell zones these estimates were combined via an arithmetic average to obtain the kriged block (zone) estimate. The covariances of the estimation errors between multiple cell zones or between single cell and multiple-cell zones were calculated by post-processing the kriging data (including the kriging weights); details of the method are described in *Zimmerman et al.*, [1990].

Transmissivity realizations were discarded when the flow simulation failed to meet a minimum performance level defined by a scalar measure of the difference between model predicted and measured hydraulic heads. The model predicted heads were interpolated (using an area-weighted interpolation method) from the grid points onto the observation points in order to calculate the head prediction errors. The mean absolute difference between model predicted and measured heads was computed and compared with a reference value; if the model's performance measure exceeded the reference value, the travel-time calculations would not be performed. The reference value was assigned as the performance measure for the head solution corresponding to the unique homogeneous transmissivity distribution which would produce the same total head drop across the system as that which is estimated from field measurements. The rationale behind this approach is that a homogeneous transmissivity distribution represents the simplest approximation to the actual transmissivity distribution (which is believed to be heterogeneous); any discretization of this homogeneous system into zones with different transmissivities should produce heads which more closely match the observed values than the simple homogeneous case.

The computer code, LHS, described in *Iman and Shortencarrier*, [1984] was used to obtain correlated samples of the transmissivity field. Both the mean vector of parameter estimates and the correlation matrix of the zoned variables were input to the code. The statistics of the sample realizations output from LHS (i.e., means, variances and covariances) matched the input statistics very well. The number of samples required for each technique was determined by observing the convergence of the first and second moments of the ground water travel-time distribution.

DESCRIPTION OF PARAMETER ESTIMATION TECHNIQUES APPLIED

Trial-and-error-inverse

The U.S. Geological Survey's hand calibrated model of the Avra Valley was used to specify the mean values of the transmissivities. Because no covariance information is derived from this technique, all the transmissivity variables were treated as uncorrelated variables by the Latin Hypercube sampling code. The variance of each transmissivity zone was assigned the variance of all the calibrated transmissivities taken over the whole field.

Kriging via Generalized Covariances

The computer code AKRIP (*Kafritsas and Bras*, [1984]) was used to estimate the covariance structure of the data and krige the field using Intrinsic Random Field theory. The Generalized Covariance model, $K(\xi)$, which best described the data was an IRF-1 of the form $K(\xi) = A\,|\xi|$ where A is a constant and ξ is the separation.

Cokriging multiple attributes

Measurements of log-transmissivity (computed from aquifer tests) and of log-specific capacity were cokriged using a computer code developed at the University of Arizona.The cross-covariance relationships were developed using the autocovariances of pseudo-variables defined as the summed pairs of the individual attributes. It can be shown that the cross-covariance can be defined in terms of the autocovariance of the summed-pair attribute and the individual attributes. We found that the cross-covariance behavior was more easily identified when the attributes were normalized to the same scale (i.e., to avoid adding attribute values that differ by orders of magnitude for the summed-pair attribute).

Cokriging for another case in which a third attribute, fractional fines, a soil characteristic, was also performed. Every transmissivity realization produced for this case resulted in computed hydraulic heads above the land surface, therefore, travel time calculations were not performed for this case. The reason for this result is that low transmissivities were estimated in the northern area of the model where transmissivity data are sparse but fractional fines data are plentiful. The fractional fines data are relatively high in the northern area, leading to lower transmissivity values because of the inverse relationship between transmissivity and fractional fines. This may indicate that vertically averaged fractional fines data may not correlate well with transmissivity data that represent the permeability of coarse-grained layers. Furthermore, the covariance relationships for the three-attribute case were developed using data in the central and southern portions of the basin; it is possible that different covariance relationships may apply in the northern region (which would violate the stationarity assumption used with this technique). This would explain the paradoxical result of obtaining a poorer model performance with the use of additional hydrogeologic information.

Ordinary kriging combined with linear regression

Log-specific capacity measurements were transformed to estimates of log-transmissivity through a regression equation developed from points at which measurements of both attributes were made. The variogram of the combined set of transmissivity estimates was used to krige the transmissivity field. Although more confidence would normally be placed on transmissivity values derived from aquifer tests, both these and the specific capacity-derived transmissivities were assigned equal weight because the kriging code did not include the option of a spatially varying nugget term.

COMPARISON OF THE PARAMETER ESTIMATES AND THEIR UNCERTAINTIES

All of the contoured parameter estimates showed very similar patterns throughout most of the model region; only in the northern area were there obvious differences. The trial-and-error estimates show a decreasing trend in the parameter values in the northern area of the model; the Generalized Covariance (GC) estimates, on the other hand, show an increasing trend in this area. The GC estimates probably continue to increase because the local trend model is, due to a paucity of data, extrapolated beyond the range of the assumed structural model. The trial-and-error estimates, while conditioned on heads, are lower than nearby measured values indicating a lack of conditioning on measured transmissivities.

Contour maps of the parameter estimation errors also showed similar patterns throughout most of the modeled region; all show lower uncertainty in the central portion of the region where most of the data lie and higher uncertainty near the margins where data are sparse.

The use of additional information (specific capacities) with cokriging causes a significant reduction in the estimation errors in the northern region of the model. The ordinary kriging combined with linear regression method resulted in even smaller estimation errors in this region; this is probably because the regression model of $log(T)$ from $log(Spc)$ is implicitly assuming a direct one-to-one correlation between $log(T)$ and $log(Spc)$, whereas this is not the case when cokriging these attributes. The estimation errors for the three-attribute case were the smallest of all, even though this case resulted in very poor model performance.

COMPARISON OF THE MODEL PREDICTION UNCERTAINTIES

The CCDF's of ground-water travel time for each technique are plotted in Figure 2. There is a marked difference between the median travel time for the trial-and-error inverse and the kriging techniques. The trial-and-error method conditions primarily on heads, not transmissivities, and over-predicts the transmissivity values in the central portion of the aquifer, leading to shorter travel times. The kriging techniques reproduce the measured values in this region which yield longer travel times. Measurements of transmissivity were made from boreholes which only penetrated the top portion of the aquifer; if these values were converted to hydraulic conductivity and multiplied by the actual aquifer thickness, the transmissivity values would be greater. Thus, the kriging techniques do not allow the full quantity of flow to pass through this portion of the aquifer because of the data that are input.

The uncertainty in the predicted ground-water travel times (i.e., the spread in the curves at the 50 percent exceedence probability level) is more than 50 percent of the shortest median ground-water travel time, indicating that significant uncertainty exists based soley on the choice of parameter estimation technique used.

The manner in which the curves for the kriging techniques are spread indicates that there is more uncertainty associated with early travel time predictions than long travel time predictions. This is probably due to the correlation structure imposed by these techniques. Because parameters are correlated, the presence of a high value of transmissivity will persist over some distance and contribute to a shorter travel time. The presence of just one block of low transmissivity along the travel path will significantly increase the travel time; the variability associated with long travel times is less because, whether it's one block or many blocks of lower transmissivity along the travel path, the result is a long travel time.

The spread within each CCDF curve shown in Figure 2 indicates that the kriging techniques have more uncertainty associated with them than does the trial-and-error inverse technique. This result is inconsistent with the statistical parameters used to control the generation of the realizations: For example, the mean estimation error assigned to each variable (for all model runs) was 0.47 for the trial-and-error inverse method. The average of the estimation errors (taken over the whole field) for all of the Generalized Covariance model runs was 0.32. Thus, one would expect the CCDF of the trial-and-error inverse to exhibit a greater variance in ground-water travel times. The reason it doesn't is because the average of the estimation errors *of the actual transmissivity fields used to compute the ground-water travel times* was 0.57 for the Generalized Covariance technique. That is, the subset of transmissivity realizations used for travel time calculations had a higher variance than the ensemble of all transmissivity realizations generated. This implies that the "good realizations" (transmissivity realizations which resulted in a satisfactory performance for the flow model) must have generally consisted of values that are "out on the tails of the parameter distributions," i.e., far from the mean estimate. This also indicates that the kriged estimates are not good indicators or not representative of the transmissivities required to reproduce the heads with this conceptual model.

The unconstrained case

The CCDF for the unconstrained case shown in Figure 2 corresponds to the travel times for the Generalized Covariance technique in which all realizations were used, regardless of how poorly the predicted heads matched the observed heads. For example, many of the discarded realizations produced heads that were hundreds to thousands of feet above the land surface. The higher median travel time predicted for this case shows that a non-conservative prediction can occur when no attempt is made to constrain the calculations to realistic realizations of input and output. For example, the probability of ground-water travel times exceeding 5,000 and 10,000 years was 12% and 0% for all cases except the unconstrained case, where these probabilities were estimated to be 64% and 24%, respectively. The distribution of travel times for this case is lognormal (rather than approximately normal as it is for the other cases), skewed toward long travel times.

SUMMARY AND CONCLUSIONS

Three kriging techniques and one inverse technique were used to obtain transmissivity estimates and estimation errors for input to a steady-state ground-water flow model of the Avra Valley aquifer, Arizona. Transmissivity realizations were generated using the Latin Hypercube sampling technique; for the kriging techniques, variable correlations were well preserved. Monte Carlo flow simulations were performed using these sample realizations and ground-water travel time calculations were performed using a particle tracking technique. CCDF's of the ground-water travel time were computed for only those runs in which the model's performance measure passed a specified performance level criterion. A CCDF for an unconstrained case, in which all runs were used regardless of the model's performance, was also computed.

The parameter estimates all showed similar patterns except in the northern area of the model were transmissivity data are lacking but head observations are available. The trial-and-error inverse technique, which conditions primarily on heads, shows a decreasing trend in transmissivity in this region, resulting in calibrated values well below the measured values.

The Generalized Covariance technique produces less reliable estimates when the observed data lie outside the normal kriging neighborhood (the area within which the assumed structural model was formulated), because the estimate is made via extrapolation of a trend rather than interpolation. The three-attribute cokriging case resulted in the smallest estimation errors, but the worst model performance. Thus, we would conclude that the best kriged estimate for use in a flow model is not necessarily the one with the lowest kriging errors.

There was approximately a 570-year spread in the median travel times among the kriging techniques; the median travel time for the trial-and-error inverse technique is about two-thirds of the median travel times predicted by the kriging techniques. The kriging techniques, which did not utilize head information, preserved the measured transmissivity values. Because the measured values only reflect the transmissivity of the top portion of the aquifer, the kriging techniques under-predicted the quantity of flow that could pass through the central portion of the model area, resulting in long travel times relative to the trial-and-error inverse method. The trial-and-error method conditions more on heads than on transmissivities, and predicted higher transmissivities than the measured values in this area. Thus, ground-water travel time predictions are dependent on the technique used to estimate the input parameters and significant uncertainty in these predictions may exist based solely on the choice of technique.

Among the kriging techniques, the CCDF curves exhibit more uncertainty at early arrival times than late times; this result is probably due to the correlation structure imposed by these techniques. The individual CCDF curves for the kriging techniques show more spread (uncertainty) than that for the trial-and-error inverse case. No variable correlations are assumed for the latter method, whereas correlations are an integral part of the kriging techniques. Thus, we observe a paradoxical result – the more correlated things are, the more uncertain the outcome.

Approximately 50% of the generated transmissivity realizations were used for travel time calculations with the trial-and-error inverse and kriging combined with linear regression methods, 20% for the Generalized Covariance method, and only 5% with the cokriging technique. The lower throughput of the latter two methods suggests that the mean estimates from these techniques are "off the mark" or inconsistent with the conceptual model and supporting data, at least in some areas of the model.

The importance of screening each simulation for realism was demonstrated with the unconstrained case. Non-conservative travel time estimates (long travel times) were obtained relative to the results from the other techniques.

REFERENCES

Iman, R. L. and M. J. Shortencarrier, 1984. "A FORTRAN 77 Program and User's Guide for Generation of Latin Hypercube and Random Samples for Use With Computer Models," Sandia National Laboratories, Albuquerque, New Mexico. NUREG/CR 3024, SAND83-2365.

Kafritsas, J. and R. L. Bras, 1984. "The Practice of Kriging (2nd Edition)," Technical Report No. 263, Ralph M. Parsons Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Zimmerman, D. A., Hanson, R. T. and P. A. Davis, 1990. "A Comparison of Parameter Estimation and Sensitivity Analysis Techniques and Their Impact on the Uncertainty in Ground Water Flow Model Predictions," Sandia National Laboratories, Albuquerque, New Mexico.

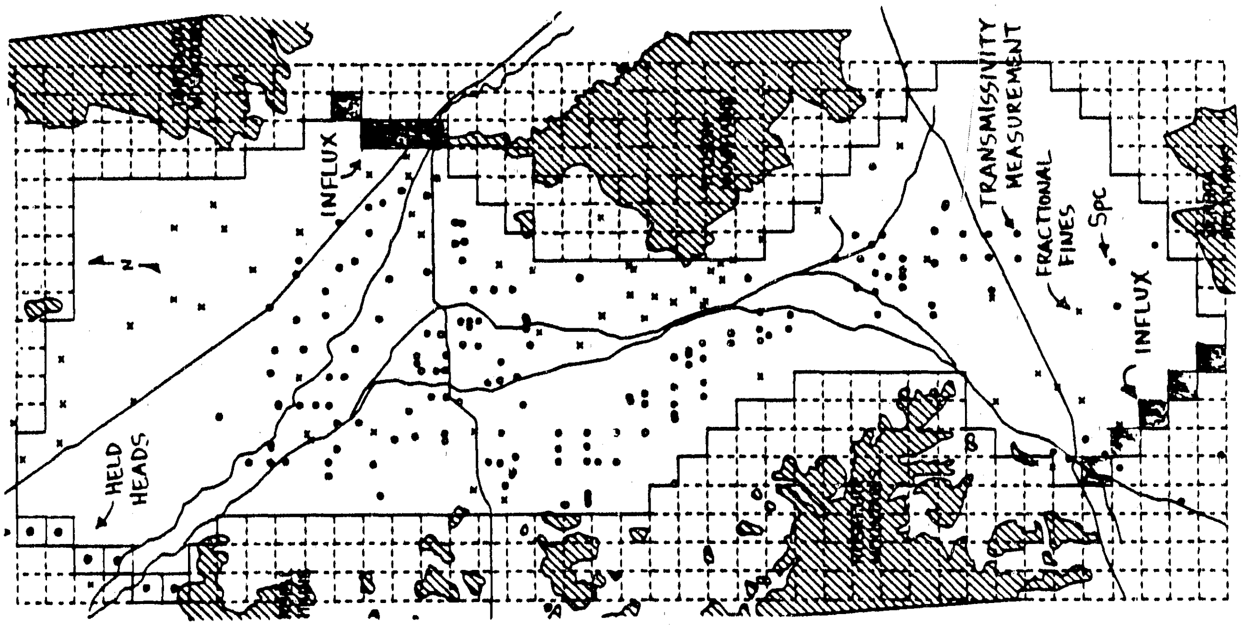

FIGURE 1. Flow domain, boundary conditions and measurement locations.

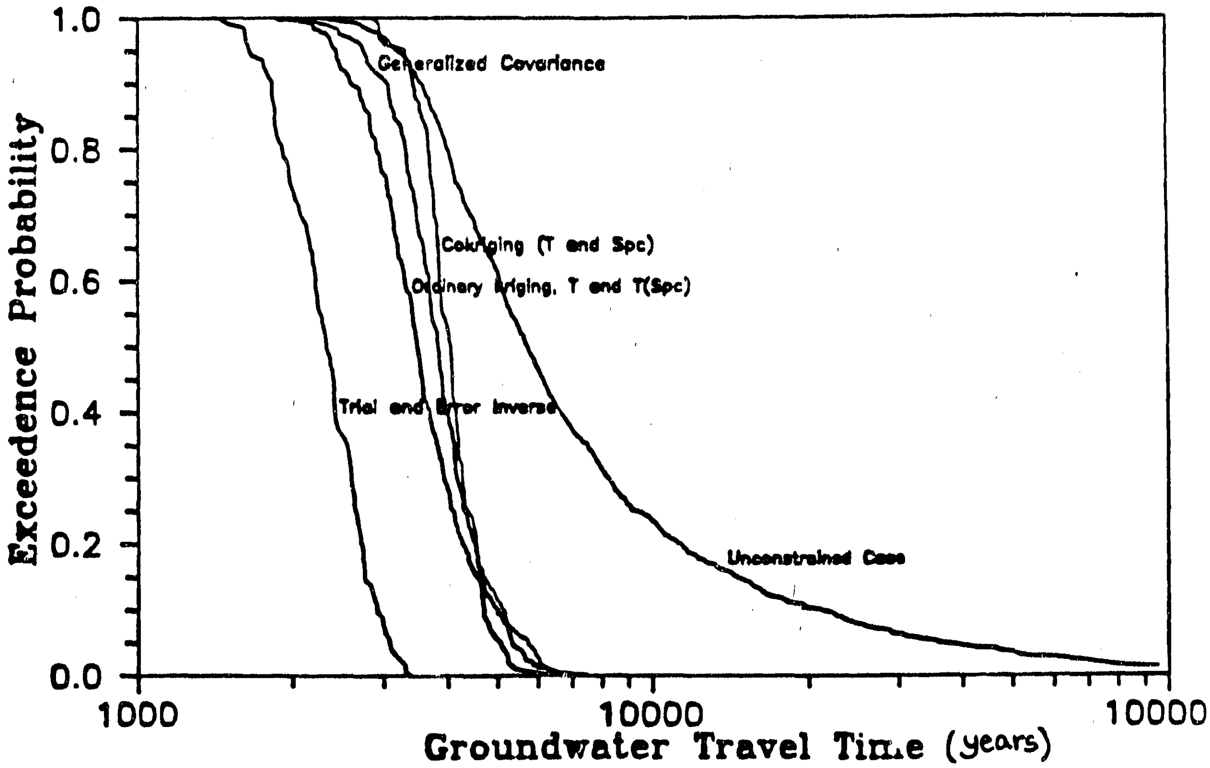

FIGURE 2. CCDF curves of groundwater travel times.

END

DATE FILMED

01/08/91