A Structured Approach for Calibrating Steady-State Ground-Water Flow Models

by T.-C. Jim Yeh and Peter A. Mock[a]

Abstract

A structured approach for calibrating two-dimensional, steady-state ground-water flow models is developed. The first step of the proposed approach is to replace the heterogeneous aquifer with an equivalent homogeneous one using the geometric mean of transmissivities. Then, boundary conditions are adjusted with the aid of a scattergram to reduce bias of the simulated hydraulic head distribution of the equivalent homogeneous aquifer. After the bias is removed, the differences between the simulated mean hydraulic head and the observed hydraulic head, resulting from small-scale heterogeneities, are then reduced by adjustment of local transmissivity values based on hydrological and geological information. The validity of the proposed procedure was tested for a hypothetical aquifer under idealized conditions and it was then applied to the Avra Valley in southern Arizona to demonstrate its utility in real world scenarios.

Introduction

Two-dimensional, numerical models for steady-state ground-water flow in porous media have been frequently used to predict ground-water flow and contaminant migration in aquifers (Anderson, 1979). Application of these models to the analysis of steady flow in an aquifer requires knowledge of the spatial distributions of transmissivity, boundary conditions, and recharge rates within the aquifer. Usually, measurements of transmissivities are too few to fully characterize the heterogeneity of the aquifer and often involve errors and uncertainties. Moreover, definitive information about boundary conditions and recharge rates rarely exists because of the complexity of the geology of the aquifer and the lack of reliable means to measure or estimate fluxes at boundaries, or recharge rates and their distribution. As a result, numerical models are often calibrated by adusting values of transmissivities, boundary conditions, and recharge rates so that simulated hydraulic head values are in agreement with hydraulic head measurements in spite of possible measurement errors. Such a calibaration (or inverse) process is often carried out either by an automatic optimization technique or by a manual (trial and error) approach. The automatic optimization approach commonly relies on some mathematical techniques to obtain a set of transmissivities, boundary conditions, and recharge rates that minimize the sum of squares of the differences between the simulated and observed heads. A thorough review of various methods along this line for calibrating a model or identifying parameter values can be found in Yeh (1986). This type of inverse approach is, however, generally confined to research and academic circles and suffers from inherent numerical difficulties (see Yeh, 1986). Practitioners, on the other hand, have relied heavily on the manual approach. This manual approach generally involves interpolation or extrapolation of transmissivity data, assigning boundary conditions based on hydrogeological information, and estimating recharge rates using simple mass balance calculations based on available precipitation and evaporation data. Subsequently, adjustments of transmissivity values and recharge rates over the entire aquifer are undertaken by trial and error to obtain hopefully ever-improving agreement between the simulated and observed hydraulic head values. Such a manual approach is often a struggle for lack of a systematic direction as to the path to a successful calibration.

In this paper, we present a structured procedure that can expedite the commonly tedious and time-consuming manual calibration process. The proposed procedure is simple and based on the concept associated with the effective parameter approach of stochastic analysis of aquifer heterogeneity (see Yeh, 1992; Gelhar, 1993). In the first section of this paper, we describe the procedure and discuss its theoretical basis. Then, the soundness of the proposed approach is tested for steady-state flow in a hypothetical aquifer with a generated random transmissivity field under idealized conditions. Finally, the procedure is applied to the Avra Valley in southern Arizona to demonstrate its utility in dealing with realistic field problems.

[a]Department of Hydrology and Water Resources, The University of Arizona, Tucson, Arizona 85721.

Received October 1994, revised April 1995, accepted May 1995.

Theory

The objective of the calibration of a ground-water flow model is to obtain good agreement between the simulated and observed hydraulic head distributions of an aquifer. Mean square error is an appropriate criterion for assessing this agreement, i.e.,

$$E[(\hat{H} - H)^2] \tag{1}$$

where E is the expected value, $\hat{H}$ and H are the simulated and observed hydraulic head values at any given location, respectively. The mean square error can also be written in terms of variance, var, and bias squared, B^2 (Priestley, 1981):

$$\begin{aligned} E[(\hat{H} - H)^2] &= \mathrm{var}\,[\hat{H}] + B^2 \\ &= E[(\hat{H} - E[\hat{H}])^2] + (E[\hat{H}] - H)^2 \end{aligned} \tag{2}$$

If we assume that measurements of hydraulic heads are error-free, and the hydraulic head is a spatial stochastic process (see Yeh, 1992; or Gelhar, 1993), the bias represents the deviation of the mean of the simulated hydraulic head from the observed head. The variance is, then, a measure of the deviation of the simulated hydraulic heads from its mean, resulting from spatial variability of transmissivities in the aquifer if the recharge is assumed to be perfectly known.

The above discussion leads to a simple rationale for calibrating a model. That is, if we conceptualize a heterogeneous aquifer as an equivalent homogeneous one, and use the effective transmissivity, recharge rates, and boundary conditions in a simulation model, simulated hydraulic heads from the model should mimic the mean (trend) of observed heads in the aquifer. In other words, the bias in the simulated head should be minimal; the variance of the simulated head should converge to the head variation resulting from the spatial variability of transmissivity. Based on this reasoning, a logical approach for calibrating a steady-state ground-water flow model first would require the determination of an effective transmissivity. A proper choice of the effective transmissivity for statistically homogeneous two-dimensional aquifers under steady-state uniform flow conditions is the geometric mean of transmissivities as reported in the literature (e.g., Mizell et al., 1982; Gelhar, 1986 and 1993; Desbarats and Srivastava, 1991). Next, boundary conditions and recharge rates are to be adjusted so that the bias in the simulated head becomes minimal. The impetus for this stems from the fact that mathematically, the governing mean flow equation becomes a boundary value problem if an effective transmissivity is identified. After the "correct" boundary conditions and recharge rates are selected, a numerical simulation can be conducted using a more detailed transmissivity map constructed from interpolation or extrapolation of available transmissivity data. The interpolation or extrapolation of the data can be carried out manually or by using mathematical tools (such as kriging or an inverse distance scheme). The final step is to reduce the variability around the mean by fine tuning the detailed transmissivity distribution.

In our proposed approach, a scattergram is suggested as an aid to the execution of the steps discussed above. The scattergram is simply an x-y plot of the simulated heads versus the measured heads with equal scales on each axis. If the simulated head at every location in an aquifer perfectly mimics the measured head, all the x-y pairs will fall on a 45° line across the scattergram. On the other hand, if the effective parameter approach is employed (i.e., the geometric mean is used as the effective transmissivity and boundary conditions and recharge rates are known), all the x-y pairs will scatter around the 45° line, indicating the minimal bias in the model. The scatter around the line can, thus, be attributed to the effect of spatially varying transmissivities at local scales neglected by the effective transmissivity in the simulation. However, a geometric mean with incorrect boundary conditions and recharge rates will produce the x-y pairs scattering around a line deviating from the 45° line. This deviation represents the bias in the simulated heads resulting from the incorrect boundary conditions and recharge rates. By adjusting these parameters, the bias can thus be eliminated, and the scatter of the data points about the 45° line can be reduced by modifying the transmissivity value in local zones or locations in the simulation domain.

Verification of the Procedure

To verify our proposed approach, a hypothetical, two-dimensional aquifer is created with 25×25 transmissivity blocks each with $\Delta x = \Delta y = 1600$ ft. The transmissivity (T) is assumed to be a stationary process with a lognormal distribution and to have a mean of 9.2, and a standard deviation of 1.0 in terms of lnT. The correlation structure is assumed to be exponential with an isotropic correlation length of one-half the size of a block. The statistics of the generated transmissivity field are: harmonic mean: 6,641 ft^2/d; geometric mean: 10,390 ft^2/d; variance of lnT: 0.9295; arithmetic mean: 16,710 ft^2/d; and variance of T: 4.68×10^8 ft^4/d^2. The distribution of the generated transmissivity is illustrated in Figure 1. The hypothetical aquifer is assumed to be bounded by impermeable boundaries on two opposite sides; the other two sides of the aquifer were assigned a constant head boundary condition.

With the above boundary conditions, steady-state flow without recharge was simulated for the heterogeneous aquifer, using the MODFLOW (McDonald and Harbaugh, 1988). During the first few simulations, it was found that the harmonic intercell averaging scheme of MODFLOW affected the statistics of the actual transmissivities used by MODFLOW when an

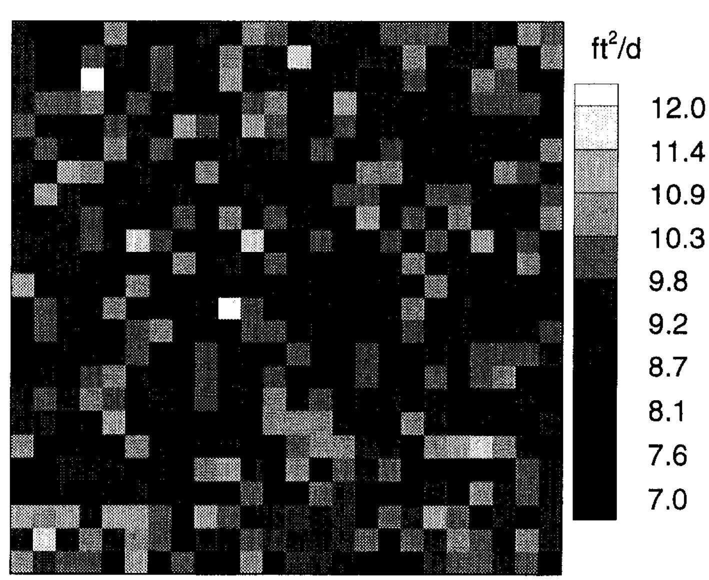

Fig. 1. Distribution of natural log of transmissivity for the hypothetical aquifer.

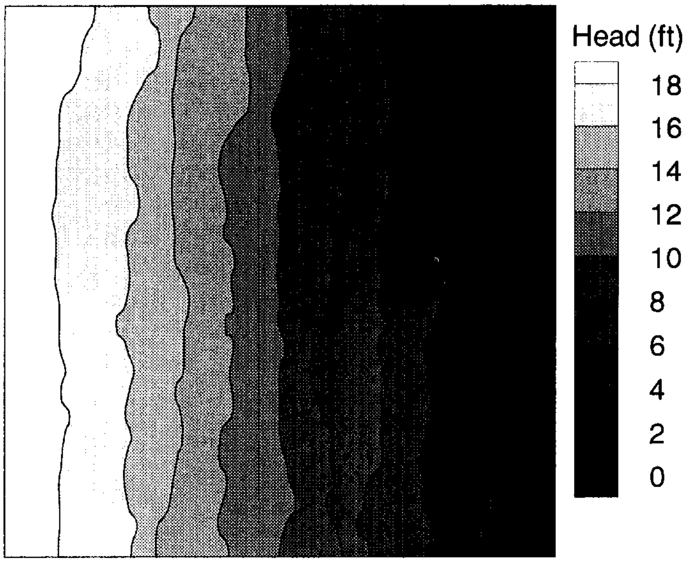

Fig. 2. Simulated hydraulic head distribution for the hypothetical aquifer.

individual finite-difference cell was assigned to each transmissivity block. This problem was reduced to a negligible level after each transmissivity block was divided into 8 × 8 small finite-difference cells. The resulting hydraulic head distribution is illustrated in Figure 2. The generated random transmissivity field and the corresponding head distribution were then used as our "real world" analog.

To illustrate the idea discussed previously, three simulation runs were conducted, where the heterogeneous transmissivity field was replaced by equivalent homogeneous aquifers with uniform values of the arithmetic, geometric, and harmonic means of the heterogeneous transmissivity field. During these simulations, one of the constant head boundary conditions was replaced with the prescribed flux boundary condition preserving the flow rates in the real world analog. In Figure 3, values of the simulated hydraulic heads of the three simulation runs are plotted against the values of the hydraulic head of the real world analog. As expected, the plot of the head distribution based on the geometric mean against that of the real world analog forms a 45° line with a scatter of 0.2738 ft^2 in variance. This variance is in good agreement with the theoretical head variance (0.1297 ft^2) based on the result by Mizell et al. (1982) for two-dimensional uniform ground-water flow:

$$\sigma_H^2 = \frac{8}{\pi^2} J^2 \lambda^2 \sigma_{\ln T}^2 \qquad (3)$$

where σ_H^2 is the head variance, J is the mean gradient (.0005), λ is the correlation scale (800 ft), and $\sigma_{\ln T}^2$ is the variance of lnT. Also shown in Figure 3 are the results based on the harmonic and the arithmetic means which are the lines above and below the 45° line, respectively, indicating bias in the simulated head. Based on these results, it is clear that selection of an incorrect effective transmissivity will produce a significant bias given that the boundary flux is known. On the other hand, if the effective transmissivity is chosen correctly, the bias must be caused by incorrect boundary conditions. Thus, our proposed procedure for modeling calibration is ratified for the two-dimensional, steady-state flow under idealized conditions.

Application to Predevelopment Avra Valley

Site Description

The Avra Valley is located just to the west of the Tucson Mountains in the Upper Sonoran Desert of Arizona. Avra Valley is oriented approximately north-south and is approximately 40 miles long (north to south) by 10 miles wide (east to west) (see Figure 4). Rainfall is approximately 12 inches per year on the basin floor, while potential evapotranspiration averages approximately 100 inches per year. The Avra Valley is underlain by as much as 9000 feet of unconsolidated to semiconsolidated sediments of primarily Quaternary age. The primary intervals of ground-water circulation are above approximately 500 feet below land surface. The reader is referred to a recent USGS study for a thorough hydrogeologic discussion of the Avra Valley (Hanson et al., 1990).

Modeling with the Proposed Approach

Several modeling studies have been conducted for the Avra Valley ground-water system during the past two decades (e.g., Moosburner, 1972; Clifton, 1981; Travers and Mock, 1984; and Hanson et al., 1990). Based on the results of these studies, predevelopment conditions in the Avra Valley aquifer can be simulated reasonably, assuming two-dimensional horizontal flow. By this, we are assuming that vertical flow components at the scale of the basin-wide simulation are negligible. This characteristic is attributed to the lack of wells pumping large volumes of water from depth during predevelopment conditions (before World War II). Another assumption made by most investigators of the Avra Valley is that natural recharge from mountain fronts and stream channels is insignificant. For a first approximation of the transmissivity distribution over such a large-scale aquifer, these assumptions appear justifiable and will be used in our study.

In this application, MODFLOW was used to simulate the head distribution. A finite-difference discretization of the Avra

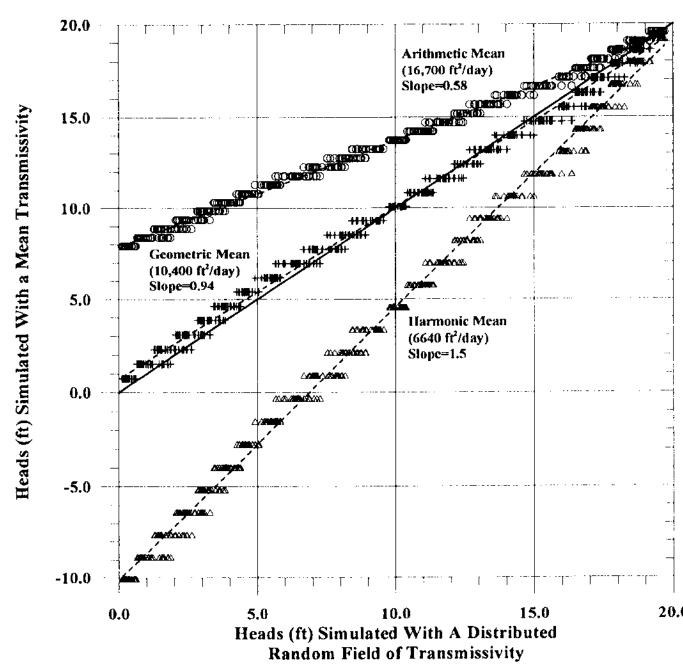

Fig. 3. Scattergram illustrating bias in the simulated heads using different mean values of the transmissivity of the hypothetical aquifer.

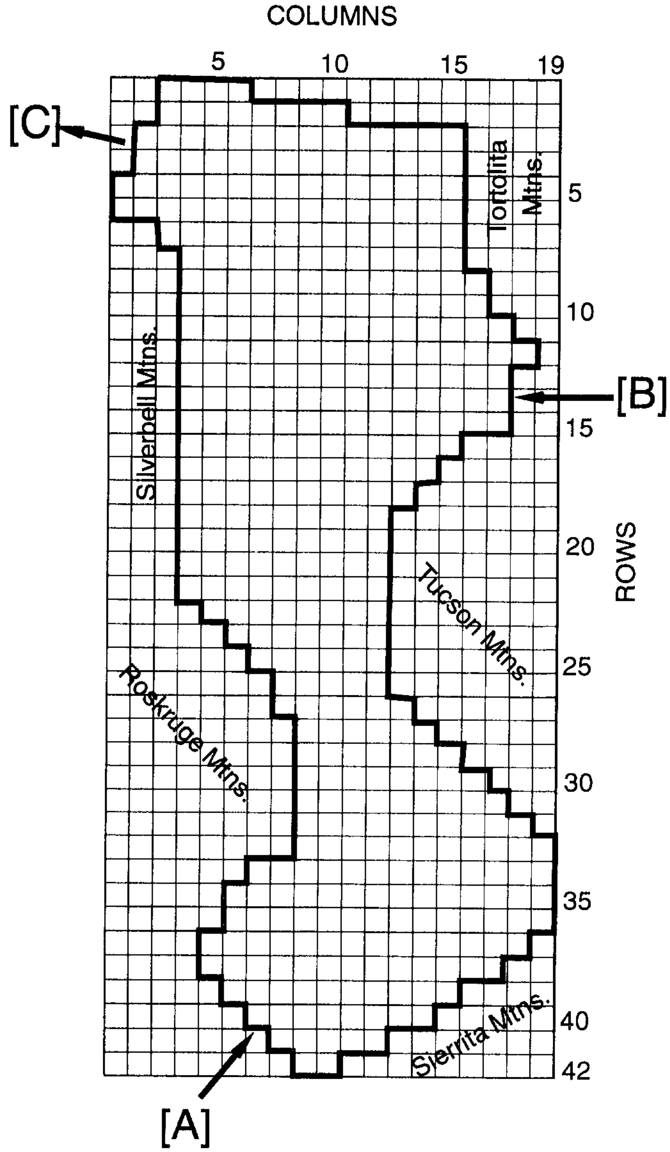

Fig. 4. Finite-difference grid of Avra Valley ground-water model.

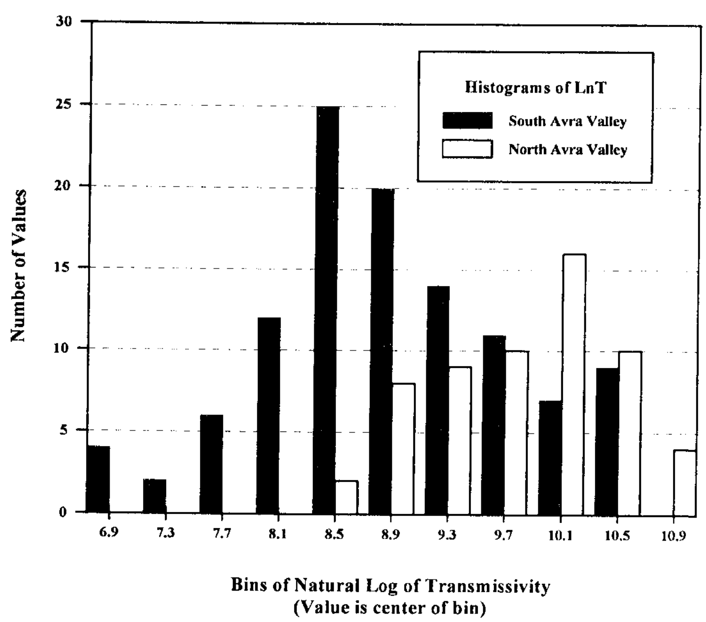

Fig. 5. Frequency distribution of natural log of transmissivity estimates in Avra Valley.

Valley and boundary conditions is shown in Figure 4. The grid of one-mile square node blocks was borrowed directly from Hanson et al. (1990) and is similar to that used by Travers and Mock (1984). The resolution appears suitable for the scale of simulation and interpreted horizontal hydraulic head gradients.

Our calibration exercise began with statistical analysis using the GEOEAS program (Englund and Sparks, 1989) of the 169 transmissivity estimates from pumping tests for the Avra Valley found in Clifton (1981). Figure 5 shows the histograms for the natural log of transmissivity distribution in Avra Valley, indicating two distinct transmissivity zones (northern and southern parts of Avra Valley, see Figure 11). There were 110 transmissivity measurements in the northern part and 59 in the southern part. These measurements were spread across each part; however, there was some concentration towards the center in each part. Ground-water flow from the Three Points inflow boundaries (A in Figure 4) crosses two different mean transmissivity zones and joins the inflow from the Rillito boundary (B in Figure 4) and discharges at the Picacho outflow boundary (C in Figure 4). After a statistical analysis, we found the geometric mean of the south data to be 7,200 ft^2/day and the geometric mean of the north data to be 17,7000 ft^2/day. This is consistent with the geologic interpretation that sediments carried by the Santa Cruz River (north zone) are from a watershed of much greater relief than sediments in the south zone.

The existence of the two geologic zones with a significant difference in the mean transmissivity value suggests that the Valley must be treated as two statistically homogeneous aquifers. We therefore started with two zones of geometric mean transmissivities, the regular grid discussed above, and constant heads at the three boundaries, based on interpretations of the predevelopment water-level map. The resulting fluxes calculated by MODFLOW at the three constant head boundaries are 1.45×10^6 ft^3/d of inflow at Rillito, 1.03×10^6 ft^3/d of inflow at Three Points, and 2.49×10^6 ft^3/d of outflow at Picacho. The simulated heads and scattergram for the simulation are illustrated in Figures 6 and 7, respectively. There clearly is a large bias and variance evident in the central part of the scattergram, and the differences between measured and simulated heads are large (the square root of MSE is 79 ft) with respect to the total head drop through the simulated system (approximately 800 ft).

The prescribed head boundary conditions at the Rillito and Three Points inflow boundaries were then converted to constant flux boundaries, and the flux values were adjusted to reduce bias as viewed on the scattergram. The outflow boundary at Picacho was maintained as a constant head boundary to ensure a unique head solution. The two geometric mean transmissivity zones (north and south) were left unchanged. Adjustments of the two inflow boundaries at Rillito and Three Points to reduce bias lead to refined estimates of: 1.02×10^6 ft^3/d of inflow at Rillito, 8.64×10^5 ft^3/d of inflow at Three Points, and 1.89×10^6 ft^3/d outflow at Picacho. These values are approximately 25% lower than previous numerical modeling studies, but we believe the new estimates are supported by the theoretical developments described before.

The scattergram for the simulated head after adjustment of fluxes is shown in Figure 8. A dramatic reduction in the bias and variance is clearly observed on the scattergram as compared with Figure 7. The differences between measured and simulated heads were significantly reduced (the square root of MSE decreased to 38 ft). These differences, as discussed previously, are a reflection

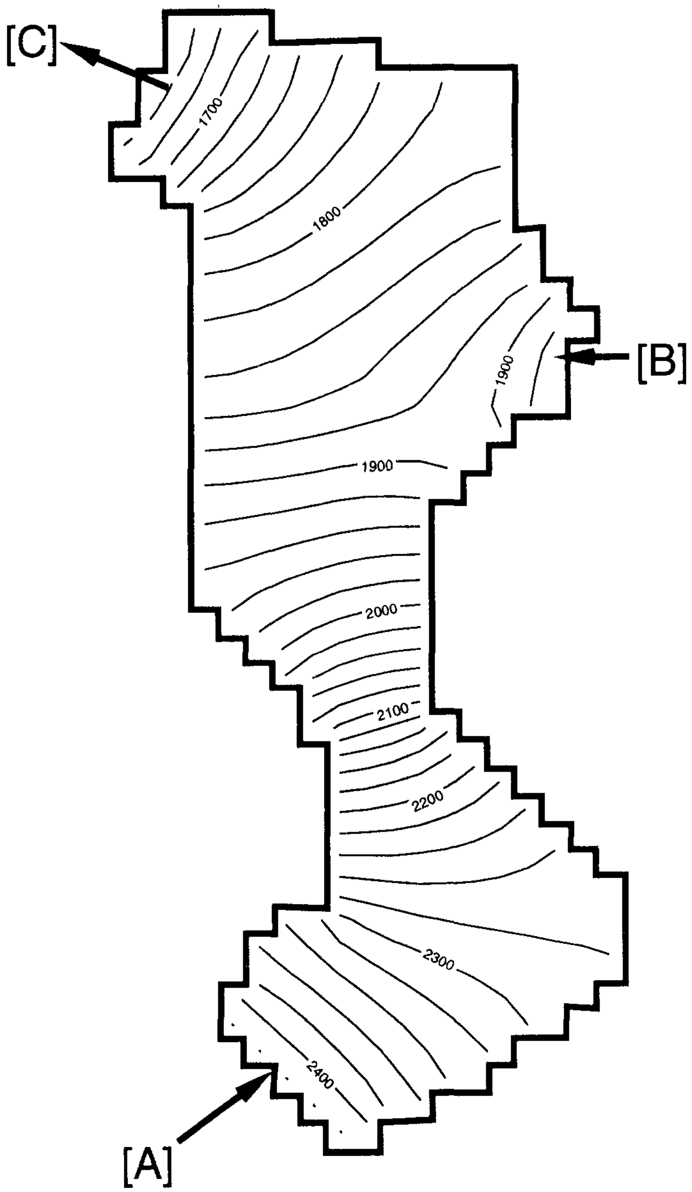

Fig. 6. Simulated hydraulic head distribution for Avra Valley, using two geometric mean transmissivity zones (A is Three Points inflow, B is Rillito inflow, and C is Picacho outflow boundaries).

of the variability in T omitted in the two transmissivity zones approach.

Our next step was to make a further reduction in the variance through the classic geologic-hydrologic, trial, and refinement method. During this step, the constant head outflow boundary condition at Picacho, and the two adjusted constant flux boundaries at Rillito and Three Points were kept unchanged. The scattergram was also used as a guide in identifying areas for refinement and a measure of the improvement (or conversely, "dead ends"—nonproductive changes in transmissivity). In effect, this procedure leads one to "unkink" each section of the scattergram, a process which can be viewed as resolving "micro-biases." We discovered also that this process of reducing local variance is most efficiently conducted by starting at one end (downstream) and working to the other (upstream), much as one would "unkink" a bent wire. That is, first one locates a kink on the scattergram and contour plot, calculates gradient changes needed, and then adjusts the value of transmissivity. Preserving

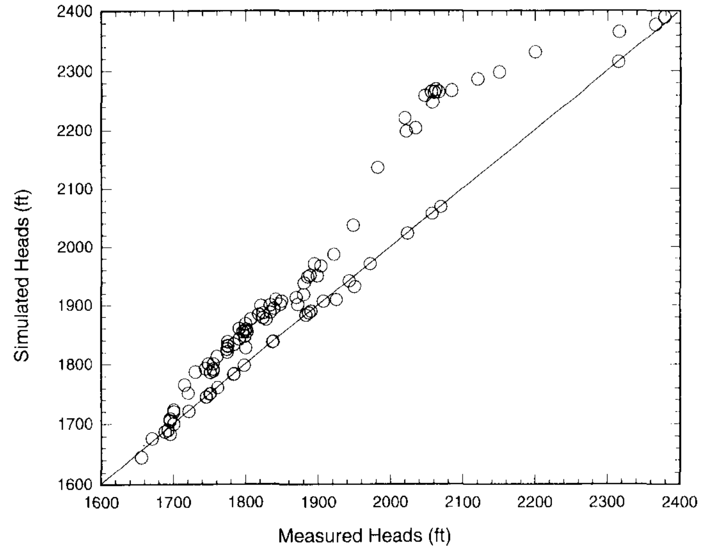

Fig. 7. Scattergram showing the bias in the simulated hydraulic head, using two geometric mean transmissivity zones without adjustment of boundary condition.

the correct overall slope maintains consistency with the geometric mean of the available point transmissivity estimates, an important advantage over typical, unstructured trial and refinement approaches. Each change to reduce variance was reviewed for geologic reasonableness with respect to the structural geology and stratigraphy of the Avra Valley's saturated sediments as best we know them.

Approximately 20 to 30 simulations and refinements to the transmissivity distribution resulted in much improvement in the simulated heads as shown in Figure 9, and in the scattergram as illustrated in Figure 10. The square root of MSE fell to 10.35 ft in the process, which could have continued further, but we felt we would have been attempting to be more precise than the water level elevations, due to the accuracy of the old topographic maps used to estimate well-head elevations. The resulting transmissiv-

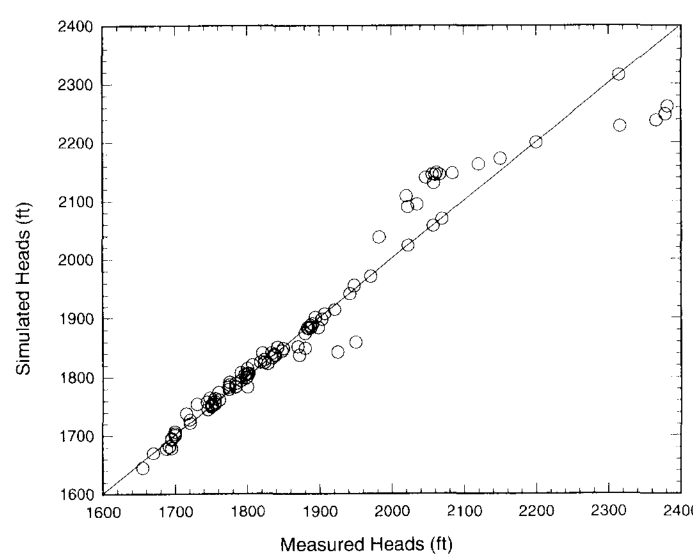

Fig. 8. Scattergram showing minimized bias after adjustment of boundary conditions.

ity distribution (Figure 11) not only provides a close fit of simulated to measured heads, but also reflects important geologic features of the area. In the south end of Avra Valley, from Three Points north for several miles, there are coalescing fans with substantial debris flows and mudslides of relatively lower transmissivity. Resorting, and therefore coarsening, of the edges of the fan materials just described occurs in the higher-energy environment of the narrows about halfway (north-south) in the basin. A zone of relatively higher transmissivity materials extends from the main drainage of the Sierrita Mountains on the southeast to the narrows. A band of very coarse unconsolidated sediments from the Upper Santa Cruz Basin (east of the Tucson Mountains) is oriented parallel to the present Santa Cruz River. Sediments transported from the northern half of the Avra Valley watershed were vigorously resorted and combined with these coarse materials from the Upper Santa Cruz Basin. Materials close to the edge of the basin are unsorted, mudslides and debris flows of relatively lower transmissivity.

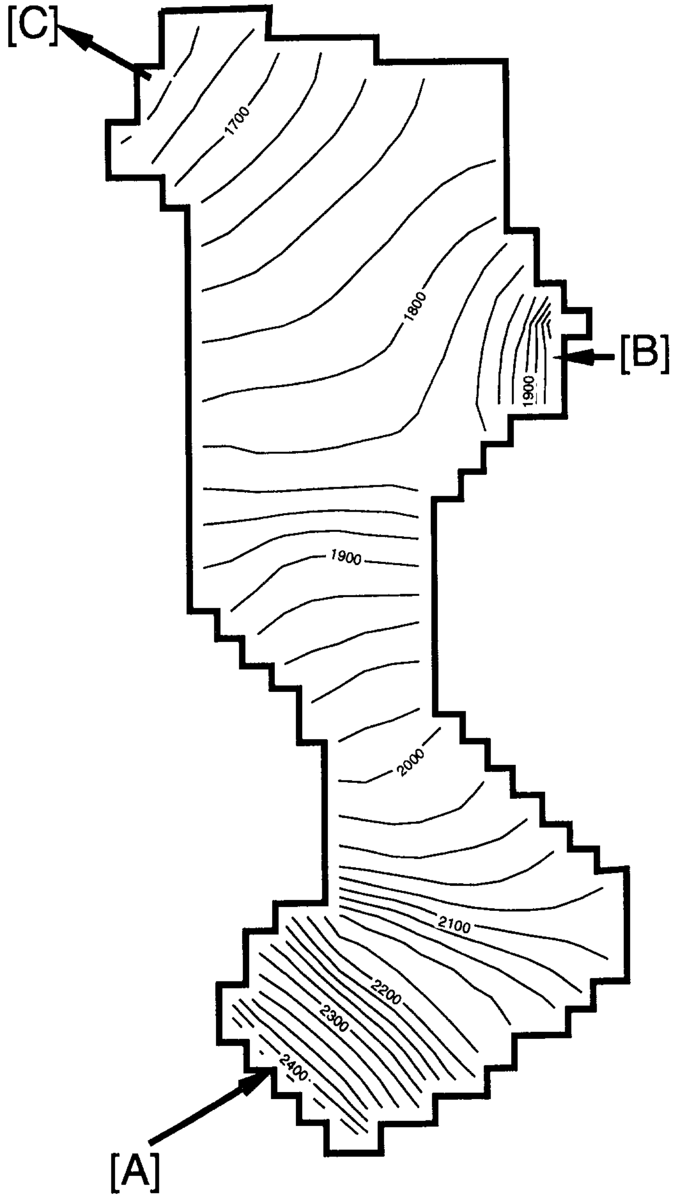

Fig. 9. Simulated hydraulic head distribution for Avra Valley, after model calibration (A is Three Points inflow, B is Rillito inflow, and C is Picacho outflow boundaries).

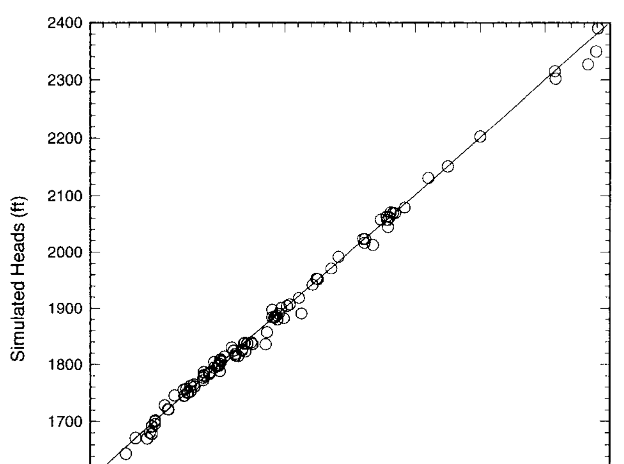

Fig. 10. Scattergram showing reduced variance after model calibration.

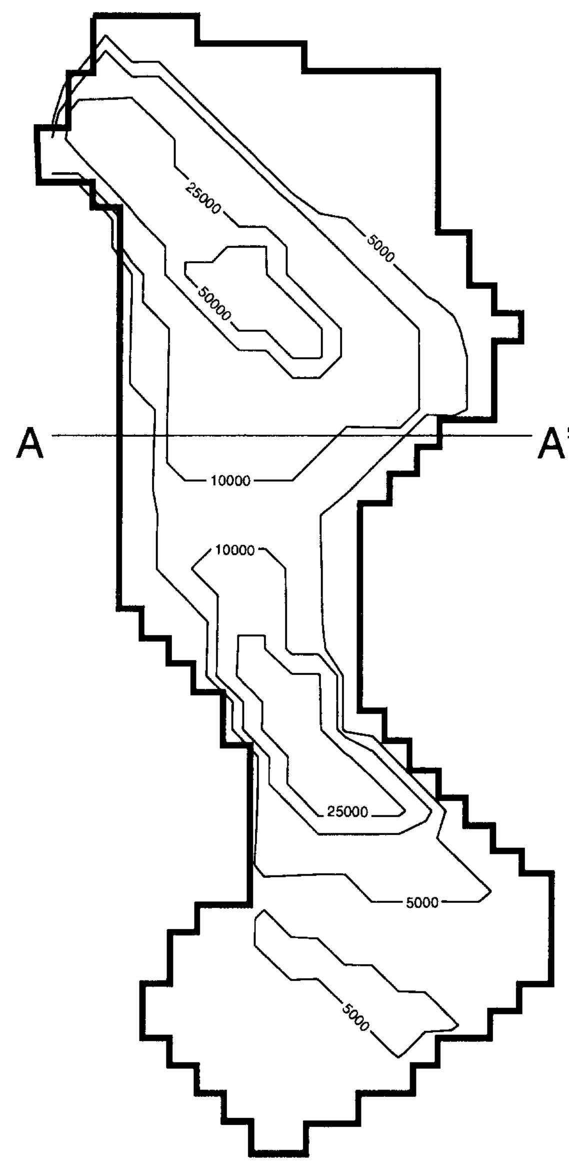

Fig. 11. Transmissivity distribution in Avra Valley, after model calibration (A-A′ is division of two geometric mean transmissivity zones).

The close fit of the simulated heads to measured heads, combined with the reasonableness of the geologic interpretation lends credence to the idea that we have a very good calibration of transmissivity for predevelopment conditions in the basin as a whole. Best of all, the calibration proceeded smoothly and logically from the downstream boundary condition to the top of each inflow area, adjusting the transmissivity to reduce local variance as we went. We estimate that without the guidance of the scattergram, over 70 simulations would have been required, and the endpoint achieved here may never have been identified conclusively.

Conclusion

The proposed simple procedure appears robust in dealing with the reproduction of the hydraulic head distribution in the Avra Valley. Various automatic procedures or different manual calibration procedures would be expected to derive a similar result. However, our proposed procedure is consistent with recently developed stochastic theory on effective parameters (Gelhar, 1993) and provides structured steps to efficiently calibrate a ground-water flow model to reduce deviations resulting from different factors at different scales. The proposed structured procedure is expected to be useful not only for the manual approach but also for automatic models. However, the manual approach based on our procedure allows modelers to easily incorporate geological information, experience, and intuitions at each step of the calibration process. Because of the logical structure of the proposed procedure, it can significantly expedite the manual calibration as compared with the unstructured one commonly used by practitioners. Finally, we would like to emphasize the fact that any calibration procedure may not necessarily lead to exact boundary conditons, transmissivity distribution, and other parameters for a field aquifer due to our incomplete knowledge of the complexity of geology, simplification of models, errors in observations, etc.

Acknowledgment

This study is a result of a class project in Advanced Ground Water Hydrology (HWR535) taught by the first author at the Department of Hydrology and Water Resources at the University of Arizona. The contributions of many students are acknowledged. Part of funding for this paper is from the NSF grant EAR-9317009.

References

Anderson, M. P. 1979. Using models to simulate the movement of contaminants through ground water flow systems. CRC Crit. Rev. Environ. Control. v. 9, pp. 97-156.

Clifton, P. M. 1981. Statistical inverse modeling and geostatistical analysis of the Avra Valley aquifer: Tucson, Arizona. Univ. of Arizona, master's thesis. 190 pp.

Desbarats, A. J. and R. M. Srivastava. 1991. Geostatistical characterization of ground water flow parameters in a simulated aquifer. Water Resour. Res. v. 27, no. 5, pp. 687-698.

Englund, E. and A. Sparks. 1989. GEO-EAS (Geostatistical Environmental Assessment Software) User's Guide. Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. EPA, Las Vegas, NV 89193-3478.

Gelhar, L. W. 1986. Stochastic subsurface hydrology from theory to applications, Water Resour. Res. v. 22, no. 9, pp. 135S-145S.

Gelhar, L. W. 1993. Stochastic Subsurface Hydrology. Prentice Hall. 390 pp.

Hanson, R. T., S. R. Anderson, and D. R. Pool. 1990. Simulation of ground-water flow and potential land subsidence, Avra Valley, Arizona. U.S.G.S. Water Resour. Invest. Rpt. 90-4178.

McDonald, M. G. and A. W. Harbaugh. 1988. A modular three-dimensional finite-difference ground-water flow model. U.S.G.S. Tech. of Water Resour. Invest. Book 6, Chap. A1, 528 pp.

Mizell, S. A., A. L. Gutjahr, and L. W. Gelhar. 1982. Stochastic analysis of spatial variability in two-dimensional steady ground water flow assuming stationary and nonstationary heads. Water Resour. Res. v. 18, no. 4, pp. 1053-1067.

Moosburner, O. 1972. Analysis of the ground water system by electrical-analog model, Avra Valley, Pima and Pinal Counties, Arizona. U.S.G.S. Hydrol. Invest. Atlas HA-215. 2 sheets.

Priestley, M. B. 1981. Spectral Analysis and Time Series. Academic Press. 890 pp.

Travers, B. C. and P. A. Mock. 1984. Groundwater modeling study of the upper Santa Cruz basin and Avra Valley in Pima and Santa Cruz Counties, southeastern Arizona. Arizona Dept. of Water Resour. unnumbered report, 2 vols., v.p.

Yeh, T.-C. J. 1992. Stochastic modeling of ground water flow and solute transport in aquifers. Hydrological Processes. v. 6, pp. 369-395.

Yeh, W. W-G. 1986. Review of parameter identification procedure in ground water hydrology: The inverse problem. Water Resour. Res. v. 22, no. 2, pp. 95-108.