Computational Methods in Water Resources XIII, Bentley et al. (eds) © 2000 Balkema, Rotterdam, ISBN 90 5809 123 6

Propagation of measurement errors in reservoir modeling

R.J. Pawar
Geoanalysis Group, Los Alamos National Laboratory, N.Mex., USA

D.M.Tartakovsky
Computer Research and Application, Los Alamos National Laboratory, N.Mex., USA

ABSTRACT: Successful modeling of flow and transport in geologic formations requires detailed characterization of their hydraulic properties. In practice, these properties are at best measured at selected locations where their values are often corrupted by experimental and interpretive errors. Such data are further used in model parameterization and subsequent numerical simulations. We investigate the propagation of measurement errors through all stages of our modeling process: (1) data collection mimicked by numerical simulations of transient pumping tests; (2) transient semilog data analysis; (3) geostatistical generation of permeability distributions with nugget; and, finally, (4) flow simulations. We demonstrate how even small experimental and interpretive errors can manifest themselves in inaccurate predictions of the system states.

1 INTRODUCTION

Experimental data for hydraulic properties, such as hydraulic conductivity, transmissivity, or permeability, are inherently subject to errors. These parameters cannot be measured directly, but instead are inferred from various forms of Darcy's experiments, which range from the classical laboratory experiment of H. Darcy to field-scale pumping tests. Consequently, permeability data are corrupted by experimental, interpretive, and model errors. Experimental errors reflect inaccuracies of the instruments and experimental procedures as well as operator errors. Interpretive errors arise from a guesswork associated with determining different test parameters such as radii of influence during pumping tests and/or time it takes to reach steady-state. Model errors are caused by misinterpretation of data through use of a wrong conceptual model, for example, not taking into account effect of leakage while interpreting data for an aquifer. Some of these errors can be eliminated or reduced by collecting data under controlled conditions of the laboratory. However, recently discovered scale effects (e.g. Neuman 1994) make the use of such data for modeling field scale processes questionable.

While the importance of measurement errors for accurate prediction of fluid flow in geological formations is fully appreciated (e.g. Clifton & Neuman 1982), there are virtually no attempts to analyze effects of error propagation through a modeling process. A typical modeling process starts with assigning permeability values to nodes or cells of a numerical mesh where experimental data are not available. This can be accomplished via any of the multiple stochastic or geostatistical approaches. In these geostatistical approaches, which incorporate experimentally observed data, the measurement errors manifest themselves through a so-called nugget effect.

The nugget effect can be caused by a combination of sub-scale fluctuations as well as measurement errors. It is important to separate the effects of the two. Indeed, the small scale fluctuations have a well-defined physical meaning and should be preserved through the nugget. On the other hand, attempts should be made to completely eliminate the sources of measurement errors or take into account their effect on the resulting realizations. Clearly, there is no point of honoring inaccurate data exactly. If a decomposition of the nugget into the sub-scale and measurement error components can be accomplished, then post-simulation filtering of Marcotte (1995) can provide an efficient algorithm for conditional prediction of flow. It is only recently that a problem of simulating flow with data subject to measurement errors has received due attention in the geostatistics (Marcotte 1995, Oliver 1996).

In this paper, we examine propagation of measurement errors through the modeling process, all the way up to the flow simulation results. While we analyzed oil flow, the same analysis can be extended to groundwater flow. In Section 2 we generate synthetic pumping test data that are analyzed by means of the transient semilog model. Starting from the premise that the parameters used in this analysis (such as pumping rates and formation thickness) are subject to the experimental and interpretive errors, we derive in Section 3 expressions for mean and standard deviation of the measured permeability. The latter serves as a measurement error component of the nugget effect. Section 4 is devoted to generating geostatistical realizations of the permeability data. Permeability distributions are generated with variogram models that take into account nugget effects based on the measurement errors. These distributions are compared to the distributions generated with variogram models that have zero nugget effects. Finally, in Section 5 we compare the effect of these different permeability distributions on the flow simulation results. This comparison demonstrates clearly that even small experimental and interpretive errors may lead to significant discrepancies in estimating the production from an oil reservoir.

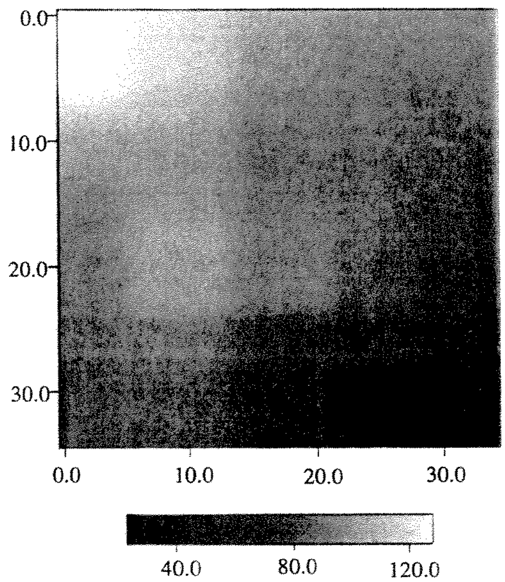


Figure 1a. Reference (true) permeability distribution.

2 PROBLEM FORMULATION

Consider a synthetically generated reservoir. The reservoir is discretized into 35 by 35 elements, and permeability values assigned to each element are considered to be our "ground truth" (Figure 1a).

This permeability field is used to generate synthetic pumping test data at 100 selected locations throughout the reservoir (Figure 1b).

These pumping tests are used to infer permeabilities through semilog well test analysis (Horne 1995), described in Section 2.2. Clearly, the "inferred" permeabilities will differ from the "true" permeabilities. In our case, this discrepancy can be attributed to numerical errors, as well as to the influence of the surrounding cells; while in actual field tests it is attributed to measurement errors.

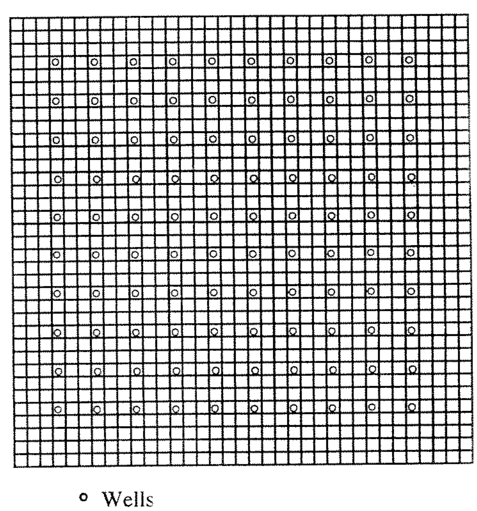


Figure 1b. A schematic of the numerical grid and locations of the wells used for data collection.

2.1 *Data acquisition*

At each of the 100 wells we numerically simulated a typical pumping test. In these simulations a grid cell with a well was subdivided into 11 by 11 sub-cells. Each sub-cell was assigned the hydraulic characteristics of the parent cell. The wells were located at the center of the corresponding parent cells.

The grid refinement was performed to better identify boundaries of the pressure disturbance due to pumping. To minimize the influence of the surrounding grid-blocks, only early pressure-time data were used for permeability calculation. In real field tests, the data collected at initial times are corrupted by the wellbore storage effects, and should be separated from the data used for calculating a slope of the pressure-time curve. This procedure is somewhat subjective and often serves as a source of interpretive errors. This source of error does not affect the data used in this study, since the synthetic well test data were generated without well bore storage.

Each well was operated at a constant pumping rate, q = 5 Stock Tank Barrel (STB) per day, and simulated pressure responses were recorded. The following parameters were used in our numerical simulations, formation thickness, h = 25 feet; oil formation volume factor, B = 1.12 Reservoir Barrels (RB) per STB, and oil viscosity μ = 4 centipoise. Figure 2 shows a data set derived from these simulations for one of the pumping wells.

2.2 *Data analysis*

Next, we employ the semilog well test analysis to interpret these synthetic pumping test data. This technique is based on the analytical solution of the transient flow equation, which gives the dependence of wellbore bottom hole pressure, P_{wf}, on time, t,

$$P_{wf} = P_i - 162.6\frac{qB\mu}{kh}\log(t) \quad (1)$$

where P_i = initial reservoir pressure, and k = permeability.

The following assumptions are essential for deriving this solution: (i) the reservoir is of infinite extent; (ii) the wellbore storage effects are neglected; and (iii) formation damage or skin effect is absent. Each of these assumptions can introduce significant interpretive errors, which are left out of the scope of the present investigation.

The synthetic well test pressure data were used to infer permeability from Equation (1). The comparison of the permeability data derived from our numerical pumping tests with the underlying, "true" permeability values reveals that the derived values are subject to errors as high as 20%.

Given the "true" and derived values of permeability, evaluating the measurement errors is trivial. In reality, however, the true values are unknown, and estimating measurement errors is often challenging. In the following section we evaluate these errors by assuming that a model used for data analysis, i.e. Equation (1), is correct and that all errors stem from imprecise measurements and their faulty interpretation.

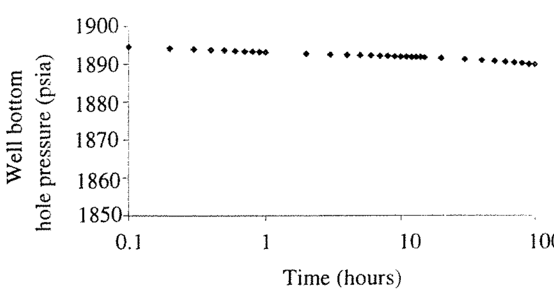


Figure 2. An example synthetic well test data.

3 ERROR ANALYSIS

The synthetic experimental data were used to calculate the slope, s, of the semilog pressure-time curve. Then permeability is readily inferred from Equation (1),

$$k = A\frac{qB\mu}{sh} \quad (2)$$

where A = 162.6. We further assume that the pumping rate, q, viscosity, μ, and the formation factor, B, are subject to experimental errors, while the slope, s, and formation thickness, h, are subject to experimental and interpretive errors. Consequently, we represent these quantities, x, as a sum of their (ensemble) means, $\overline{x}$, and zero mean random fluctuations, x' (e.g. $q = \overline{q} + q'$). In this context, $\overline{x}$ can be thought of as the true value of x, while x' represents measurement errors. The condition $\overline{x'} = 0$ indicates that the measurements are unbiased. Otherwise, the instruments should be re-calibrated, and the bias be subtracted from the mean.

3.1 *General analysis*

It follows from Equation (2) that the ensemble mean of permeability, or its "true" value, can be evaluated by taking the mean of

$$k = \frac{A(\overline{q} + q')(\overline{B} + B')(\overline{\mu} + \mu')}{(\overline{s} + s')(\overline{h} + h')} \quad (3)$$

To evaluate this mean, one has to make some assumptions regarding statistical distributions of the random errors in Equation (3). Assuming that these errors are mutually uncorrelated, we obtain the first-order (in $\sigma_x/\overline{x}$, where σ_x is the standard deviation of x) approximation of $\overline{k}$,

$$\overline{k} = A\frac{\overline{q}\overline{B}\overline{\mu}}{\overline{s}\overline{h}}\left(1 + \frac{\sigma_s^2}{\overline{s}^2}\right)\left(1 + \frac{\sigma_h^2}{\overline{h}^2}\right) \quad (4)$$

In deriving Equation (4), it was assumed that $\sigma_x/\overline{x} << 1$. Alternatively, one can easily derive an exact expression for $\overline{k}$ by specifying a particular distribution for the errors in Equation (3). We prefer to use Equation (4), since the condition $\sigma_x/\overline{x} << 1$ is satisfied in most experiments, while the data needed for deriving probability distributions are rarely available.

Measurement errors can now be characterized by permeability variance, σ_k^2,

$$\sigma_k^2 = \overline{k^2} - \left(\overline{k}\right)^2 \qquad (5)$$

where the first-order approximation of $\overline{k^2}$ is given by

$$\overline{k^2} = \frac{A^2\left(\overline{q}^2 + \sigma_q^2\right)\left(\overline{B}^2 + \sigma_B^2\right)\left(\overline{\mu}^2 + \sigma_\mu^2\right)}{\overline{s}^2\overline{h}^2} \times \left(1 + \frac{3\sigma_s^2}{\overline{s}^2}\right)\left(1 + \frac{3\sigma_h^2}{\overline{h}^2}\right) \qquad (6)$$

In field experiments, σ_q^2 is determined by the precision of the flow rate measuring device, while variances of the remaining parameters, as well as their ensemble means, can be inferred from repeated experiments. Equations (4) – (6) can be used for the error sensitivity analysis.

3.2 *Data error analysis*

In analyzing measurement errors, we assume the values of the parameters q, s, B, h, and μ, used to generate the synthetic well test data, to be identical to their corresponding ensemble mean counterparts. Standard deviations of these parameters were calculated as a percentage of the corresponding means. Parameters h and s are often corrupted by interpretive errors, while parameters q, B, and μ are subject to smaller experimental errors. Consequently, standard deviations of the first group of parameters were taken to be 1-10% of the corresponding means; while 1-5% range was chosen for the second group. Note that all these random fields are statistically inhomogeneous.

These values were subsequently used in Equations (4) – (6) to evaluate $\overline{k}$ and σ_k^2 for each of the 100 grid blocks where the measurements are available. Table 1 shows how experimental and interpretive errors (Columns I and II, respectively) get magnified during the permeability estimation (Column III). The values of experimental and interpretive errors are tabulated as the coefficient of variation, $\sigma_x/\overline{x}$, where x = s, h for interpretive errors and q, B, μ for experimental errors and the corresponding permeability coefficient of variation, $\sigma_k/\overline{k}$. Column IV shows the corresponding nugget value, which, in the absence of sub-scale fluctuations, is given by the variance of the maximum permeability measurement errors for the 100 data points.

Table 1. Coefficients of variation for the experimental (Column I), interpretive (Column II), and inferred permeability (Column III) errors and the corresponding nugget (Column IV).

I	II	III	IV
0.01	0.01	0.022	0.140
0.02	0.01	0.037	0.390
0.05	0.01	0.088	2.160
0.01	0.02	0.033	0.310
0.02	0.02	0.044	0.560
0.05	0.02	0.091	2.340
0.01	0.05	0.073	1.490
0.02	0.05	0.079	1.750
0.05	0.05	0.111	3.540
0.01	0.10	0.140	5.770
0.05	0.10	0.165	7.920

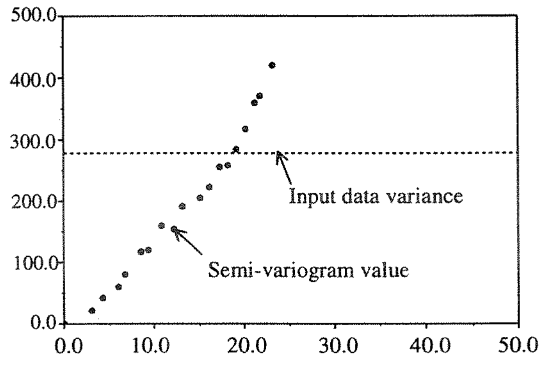


Figure 3. Semi-variogram for permeability data.

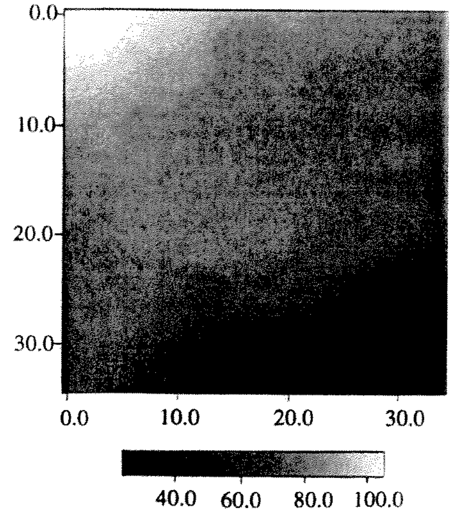
One can see from Table 1 that even relatively small experimental and interpretive errors give rise to significant errors in the inferred permeability values.

4 GEOSTATISTICAL ANALYSIS

The next stage in the propagation of the measurement errors consists of estimating unknown values of permeability through geostatistical analysis. The first step in this process is to obtain the spatial correlation parameters from a semivariogram. Figure 3 depicts such a semi-variogram for 100 data points.

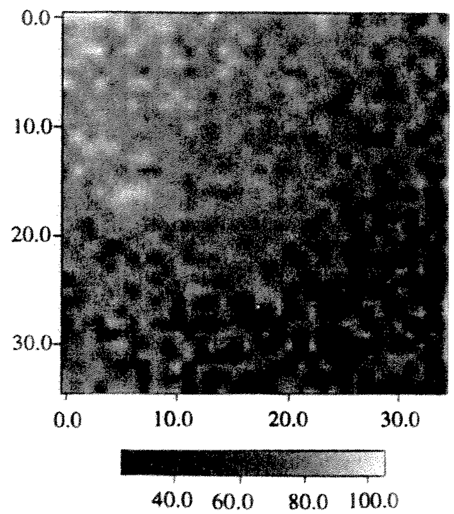
Fitting a power-law model to this variogram yielded correlation length equal to 18 grid block units, sill – 277, and power law exponent - 0.8. We used these spatial correlation parameters to stochastically generate permeability realizations using the conditional sequential Gaussian simulation technique. The effect of measurement errors was studied by changing the value of the nugget effect in the variogram model. The nugget, or variance of the measurement errors, was taken from Table 1.

The resulting permeability distributions are shown in Figures 4 and 5, for nugget zero (no measurement error) and 7, respectively.



Permeability (mili darcy)

Figure 4. Permeability distribution resulted from zero nugget.



Permeability (mili darcy)

Figure 5. Permeability distribution resulted from nugget = 7.

These distributions were generated using the same random path and conditioned to the same permeability data. One can see that the nugget effect adds variability to the resulting distributions. We further consider the effect of these realizations on fluid flow.

5 FLOW SIMULATIONS

Flow simulation models were developed with the geostatistically generated permeability fields. All of the properties, except the permeability distribution, were kept the same and the resulting oil production predictions were compared. The results for two dif-

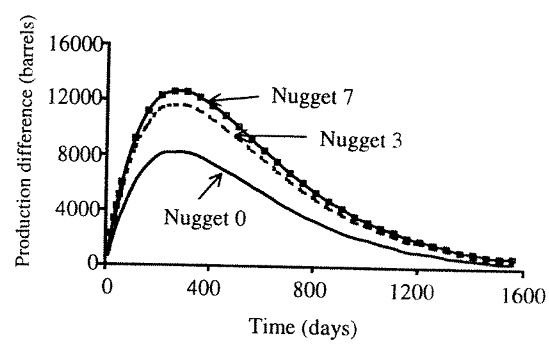


Figure 6. Absolute differences for a single well production corresponding to nuggets 0, 3 and 7.

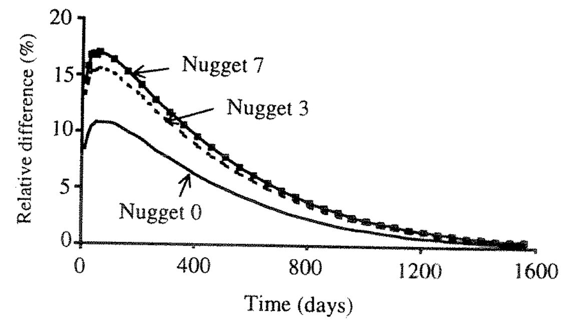


Figure 7. Relative differences for a single well production corresponding to nuggets 0, 3 and 7.

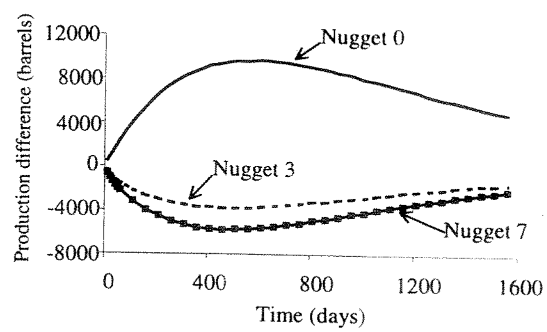


Figure 8. Absolute differences for a single well production corresponding to nuggets 0, 3 and 7.

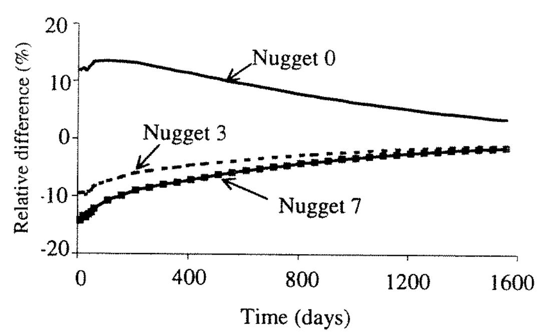


Figure 9. Relative differences for a single well production corresponding to nuggets 0, 3 and 7.

ferent wells are shown in Figures 6 – 9. The Figures show the differences in the cumulative oil production for the permeability fields generated with nuggets 0, 3 and 7, and the "true" permeability field shown in Figure 1a. Figures 6 and 8 show the absolute difference, while Figures 7 and 9 show the difference as a percentage of the prediction for the "true" permeability field.
As can be seen from the Figures, the oil productions differ significantly from their true value. The permeability realizations can both over-predict or under-predict the production response from a well. The difference (either positive or negative) increases with the nugget value. It should be remembered that the nugget in these realizations represents the variance in the measurement errors in the permeability values.

6 CONCLUSIONS

Our paper leads to the following major conclusions.

1. Experimental and interpretive procedures involved in inferring permeability from the pumping test data introduce errors in the calculated values. Permeability measurement errors depend on the input parameter errors in a nontrivial manner. We derive this dependence for the semilog well test analysis model.
2. The ultimate errors can be significant even for relatively small values of the measuring and interpretive errors.
3. Measurement errors propagate further through the modeling process as a nugget in semivariogram. The latter was used for geostatistical reconstruction of permeability fields.
4. Flow response of these geostatistically generated permeability fields can be significantly different from the "true" response or even from a geostatistical model with zero nugget effect.

REFERENCES

Clifton, P.M. & S.P. Neuman 1982. Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in southern Arizona. *Water Resour. Res.* 18(4): 1215-1234.

Horne, R.N. 1995. *Modern well test analysis*. Palo Alto: Petroway Inc.

Marcott, D. 1995. Conditional simulation with data subject to measurement error: post-simulation filtering with modified factorial kriging, *Math. Geology*. 27(6): 749-762.

Neuman, S.P. 1994. Generalized scaling of permeabilities: validation and effect of support scale. *Geophys. Res. Lett.* 21(5): 349-352.

Oliver, D.S. 1996. On conditional simulation to inaccurate data. *Math. Geology*. 28(6): 811-817.

PROCEEDINGS OF THE XIII INTERNATIONAL CONFERENCE ON COMPUTATIONAL METHODS IN WATER RESOURCES / CALGARY / ALBERTA / CANADA / 25 – 29 JUNE 2000

Computational Methods in Water Resources

Edited by

L.R. Bentley
University of Calgary, Alb., Canada

J.F. Sykes
University of Waterloo, Ont., Canada

C.A. Brebbia
Wessex Institute of Technology, UK

W.G. Gray
University of Notre Dame, South Bend, Ind., USA

G.F. Pinder
University of Vermont, Burlington, Vt., USA

OFFPRINT

A.A. BALKEMA / ROTTERDAM / BROOKFIELD / 2000