WATER RESOURCES RESEARCH, VOL. 37, NO. 12, PAGES 3019–3041, DECEMBER 2001

# Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff 2. Equivalent parameters, high-resolution stochastic imaging and scale effects

Velimir V. Vesselinov,[1] Shlomo P. Neuman, and Walter A. Illman[2]
Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA

**Abstract.** In paper 1 of this two-part series we described a three-dimensional numerical inverse model for the interpretation of cross-hole pneumatic tests in unsaturated fractured tuffs at the Apache Leap Research Site (ALRS) near Superior, Arizona. Our model is designed to analyze these data in two ways: (1) by considering pressure records from individual borehole monitoring intervals one at a time, while treating the rock as being spatially uniform, and (2) by considering pressure records from multiple tests and borehole monitoring intervals simultaneously, while treating the rock as being randomly heterogeneous. The first approach yields a series of equivalent air permeabilities and air-filled porosities for rock volumes having length scales ranging from meters to tens of meters, represented nominally by radius vectors extending from injection to monitoring intervals. The second approach yields a high-resolution geostatistical estimate of how air permeability and air-filled porosity, defined on grid blocks having a length scale of 1 m, vary spatially throughout the tested rock volume. It amounts to three-dimensional pneumatic "tomography" or stochastic imaging of the rock. Paper 1 described the field data, the model, and the effect of boreholes on pressure propagation through the rock. This second paper implements our inverse model on pressure data from five cross-hole tests at ALRS. We compare our cross-hole test interpretations by means of the two approaches with earlier interpretations by means of type curves and with geostatistical interpretations of single-hole test data. The comparisons show internal consistency between all pneumatic test interpretations and reveal a very pronounced scale effect in permeability and porosity at ALRS.

## 1. Introduction

*Vesselinov et al.* [this issue] (hereinafter referred to as paper 1) described a three-dimensional numeric inverse model for the interpretation of cross-hole pneumatic tests in unsaturated fractured tuff at the Apache Leap Research Site (ALRS) near Superior, Arizona. The model incorporates the automatic grid generator X3D [*Trease et al.*, 1996], the finite volume simulator FEHM [*Zyvoloski et al.*, 1997], a parallelized version of the parameter estimator PEST [*Doherty et al.*, 1994], and the geostatistical package GSTAT [*Pebesma and Wesseling*, 1998]. It simulates airflow on a three-dimensional grid of structured and unstructured tetrahedral elements, which represents quite accurately the geometry of vertical and inclined boreholes at ALRS.

The model simulates single-phase airflow through a porous continuum. On the basis of earlier findings by our group [*Guzman et al.*, 1994, 1996; *Guzman and Neuman*, 1996; *Chen et al.*, 2000; *Illman and Neuman*, 2001, 2000] we take this continuum to represent primarily interconnected fractures at the site. The FEHM simulator solves the airflow equations in their original nonlinear form and accounts directly for the ability of all packed-off borehole intervals to store and conduct air through the system. In paper 1 we showed that packed-off borehole test intervals have a considerable effect on airflow and that disregarding this effect may cause a systematic bias in the estimation of pneumatic parameters.

Our model treats log air permeability $\log_{10} k$ and log air-filled porosity $\log_{10} \phi$ either as being uniform throughout the rock volume or as forming random fractal fields of the fractional Brownian motion (fBm) type. In the first case, the estimated parameters represent equivalent values over the rock volume. In the second case, they describe the spatial variation of local pneumatic properties throughout this volume. Our model characterizes this spatial variability by a power variogram and estimates it geostatistically by kriging on the basis of a set of pilot points. The estimation entails simultaneous inversion of pressure records from multiple observation intervals and cross-hole tests. It thus amounts to high-resolution pneumatic tomography (or stochastic imaging) of the rock, an idea originally proposed in the context of hydraulic cross-hole tests by *Neuman* [1987].

Paper 1 included a description of five cross-hole tests labeled PP4, PP5, PP6, PP7, and PP8. Here we focus on their interpretations by means of the two approaches just described and compare them with earlier interpretations by means of type

[1]Now at Hydrology, Geochemistry, and Geology Group, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.
[2]Now at Department of Geosciences, University of Iowa, Iowa City, Iowa, USA.



Paper number 2000WR000135.
0043-1397/01/2000WR000135$09.00

curves [*Illman and Neuman*, 2001], as well as geostatistical interpretations of single-hole test data [*Chen et al.*, 2000; *Vesselinov*, 2000]. Before doing so, we remind the reader that whereas tests PP4, PP5, PP6, and PP7 were conducted with the packer configuration depicted in Figure 1 of paper 1, test PP8 was done with a slightly different packer arrangement. Whereas during tests PP4 and PP8 air was injected at a constant mass rate into the middle interval of borehole Y2, during tests PP5, PP6, and PP7 it was injected into the middle intervals of boreholes X2, Z3, and W3, respectively. The rates of injection varied from test to test. The 16 test boreholes contained 38 packed-off intervals that varied in length from 0.5 to 42.6 m. The length of the injection interval varied between 1 m (PP7) and 2 m (PP4, PP5, PP6, and PP8). Distances between injection and observation intervals ranged from 1 to 30 m. Most, but not all, of the test intervals have yielded reliable pressure records. These include 32 intervals during test PP4, 21 during test PP5, 23 during PP6, 29 during PP7, and 31 during PP8. We thus have a total of 136 reliable records to work with.

A preliminary inspection of the available cross-hole test data in paper 1 has provided support for our model by suggesting that air-filled fractures at ALRS are pneumatically well connected. Though some of the data appear to be influenced by discrete fractures, this influence is strictly local and the fracture network as a whole should be amenable to representation by a three-dimensional porous continuum. The data show that this fracture continuum is nevertheless strongly heterogeneous, providing support for our treatment of it as a random field (our decision to represent it as a random fractal of the fBm type is based on a geostatistical analysis of earlier single-hole test data [*Chen et al.*, 2000]). Some of the cross-hole data reveal evidence of nonlinear behavior and borehole storage effects due to the high compressibility of air, phenomena for which we account explicitly in our model. Temperatures remain relatively stable except in injection interval Y2-2 during tests PP4 and PP8. Since the effect is localized and not observed in the surrounding intervals, we feel comfortable modeling airflow during cross-hole tests at ALRS as being isothermal. We therefore expect our inverse model to reproduce with reasonable fidelity the conditions that prevailed during the above five cross-hole tests at ALRS. The extent to which it does so would be a test of our models' validity.

## 2. Filtering of Cross-Hole Pressure Records

To interpret the cross-hole tests with our inverse model, we filter the available pressure records so as to focus on signals that appear to be due primarily to air injection and to reduce the large set of recorded pressures down to a manageable number without significant loss of information. We do so by ignoring those portions of a pressure record that we deem strongly influenced by barometric pressure fluctuations or other extraneous phenomena and by representing the remaining portions via a relatively small number of "match points." We distribute the match points more or less evenly along the log-transformed time axis so as to capture with equal fidelity both rapid pressure transients at early time and more gradual pressure variations at later times. Ultimately, we characterize the pressure records of test PP4 by means of 252 match points, test PP5 by 132 points, PP6 by 105, PP7 by 155, and PP8 by 210 match points. In the inverse model we assign an equal weight of 1 to each of these points on the assumption that they are equally representative of pneumatic pressure behavior at ALRS.

## 3. Inverse Estimation of Equivalent Parameters

We start by analyzing pressure data from one monitoring interval at a time while treating the medium as if it was uniform across the site. Each such numerical inversion requires ~80 forward simulations and takes ~4 hours on the University of Arizona SGI Origin 2000 multiprocessor supercomputer. The procedure yields an estimate of log air permeability $\log_{10} k$ and log air-filled porosity $\log_{10} \phi$ for each pressure record. We consider these estimates to represent equivalent pneumatic parameters for a rock volume having a length scale of the order of the distance between the corresponding injection and monitoring intervals. As the rock is heterogeneous, the estimates differ from one pressure record to another. As different pressure records correspond to different distances between injection and monitoring intervals, the scales associated with the estimates also vary from record to record. We view the collection of all parameter estimates obtained in this manner, from the totality of pressure records, as a sample from a random population of equivalent log air permeabilities and log air-filled porosities representing nominal scales of meters to tens of meters. We expect them to be comparable to pneumatic parameters determined by *Illman and Neuman* [2001] on the basis of individual pressure records from test PP4 by means of analytically derived type-curves.

Figure 1 compares observed (dots) and computed (curves) records of pressure buildup versus time corresponding to cross-hole test PP4. Open circles indicate match points used to fit the inverse model to the data. Similar comparisons are depicted for test PP5 in Figure 2 and for test PP6 in Figure 3; those for tests PP7 and PP8 can be found in the work by *Vesselinov* [2000]. Some of the matches are very poor, some are of intermediate quality, and some are good to excellent. The inability of our inverse model to reproduce all pressure records properly stems in part from our representation of a heterogeneous rock by means of a uniform medium and in part from extraneous signals such as barometric fluctuations that our model does not attempt to reproduce.

Table 1 lists values of $\log_{10} k$ and $\log_{10} \phi$ and their respective individual 95% confidence limits based on data from test PP4. All log air permeability estimates appear to be reasonable for a fractured tuff. Three of the records (X3, Z2-3, Z2-4) yield unrealistically high air-filled porosity estimates ($\log_{10} \phi = -0.3$ or $\phi \approx 0.5$) equal to the upper limit that we allow this parameter to take. The corresponding fits between computed and observed pressures are poor. The records are associated with intervals that appear to be pneumatically connected to the atmosphere, which masks pressure transients and is thus much more detrimental to the estimation of porosities than that of permeabilities. We therefore ignore the corresponding porosity estimates in the ensuing discussion. The remaining estimates yield better matches. Summary statistics for $\log_{10} k$ and $\log_{10} \phi$ estimates for all five tests are listed in Tables 2 and 3.

It is of interest to note that the 2-m injection interval in cross-hole test PP4 (Y2-2) virtually coincides with one of the single-hole test intervals (labeled JG0921 by *Guzman et al.* [1996]). Though the injection rate during PP4 exceeded that during the single-hole test by >2 orders of magnitude, pressure records collected in the injection intervals during these two

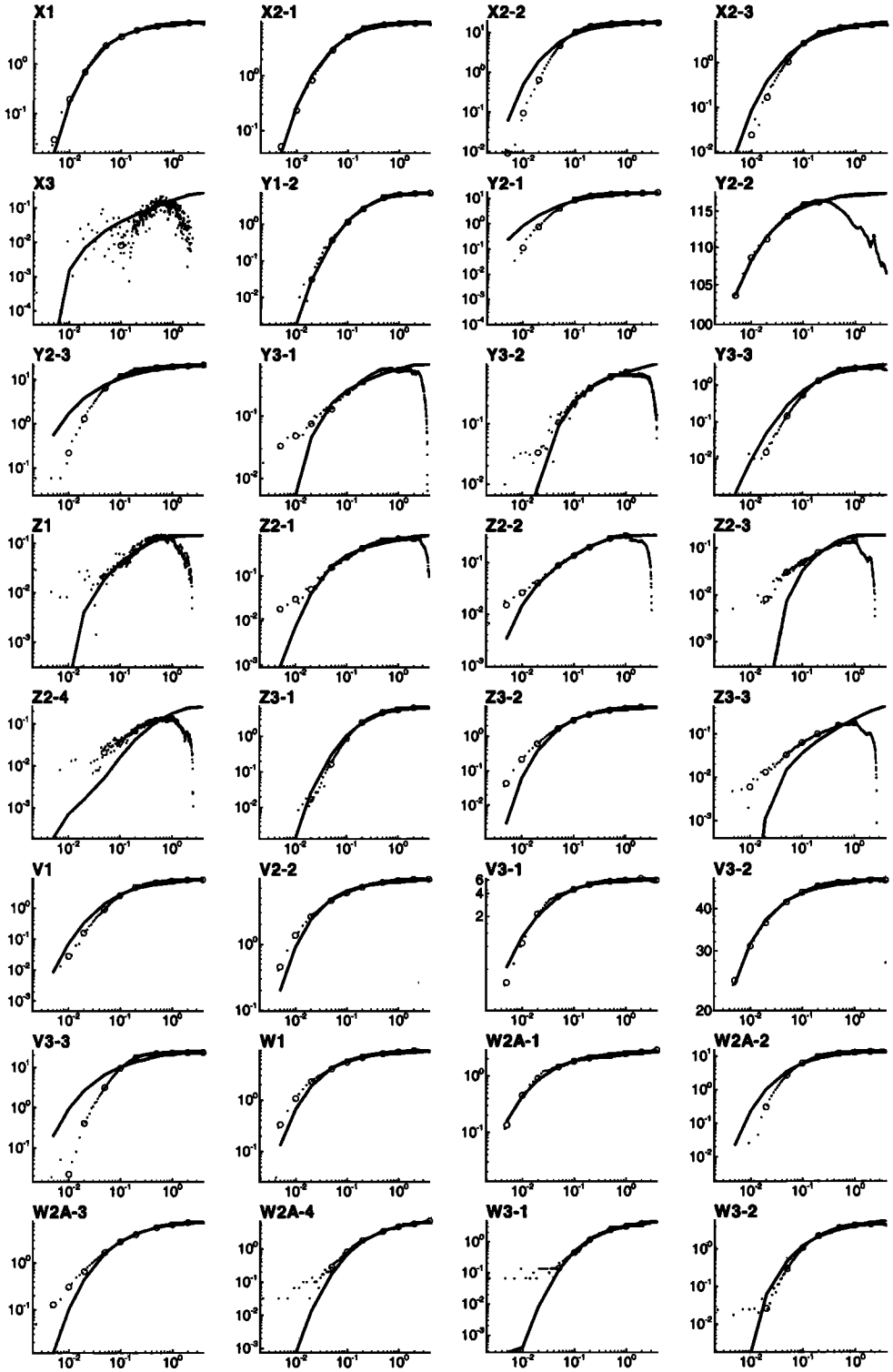


**Figure 1.** Observed (small dots) and simulated (curves, uniform rock) pressure buildup (kPa) versus time (days) during test PP4. Open circles are match points.

tests yield comparable inverse parameter estimates [*Vesselinov*, 2000]: $k = 2.1 \times 10^{-14}$ and $\phi = 6.0 \times 10^{-2}$ in the first case and $k = 2.4 \times 10^{-14}$ and $\phi = 1.4 \times 10^{-2}$ in the second case. In each case, we take these parameters to represent the equivalent properties of a small rock volume in the immediate vicinity of the injection interval, nominally equal in scale to the 2-m length of the injection interval. This is consistent with our interpretation of parameter estimates, obtained from the inversion of individual pressure records in a heterogeneous rock, as equivalent values representative of a rock volume that lies between the monitoring and the injection interval (in the case just discussed, the two intervals coincide, and so the relevant rock volume lies close to the test interval).

The above results correspond to a barometric pressure of

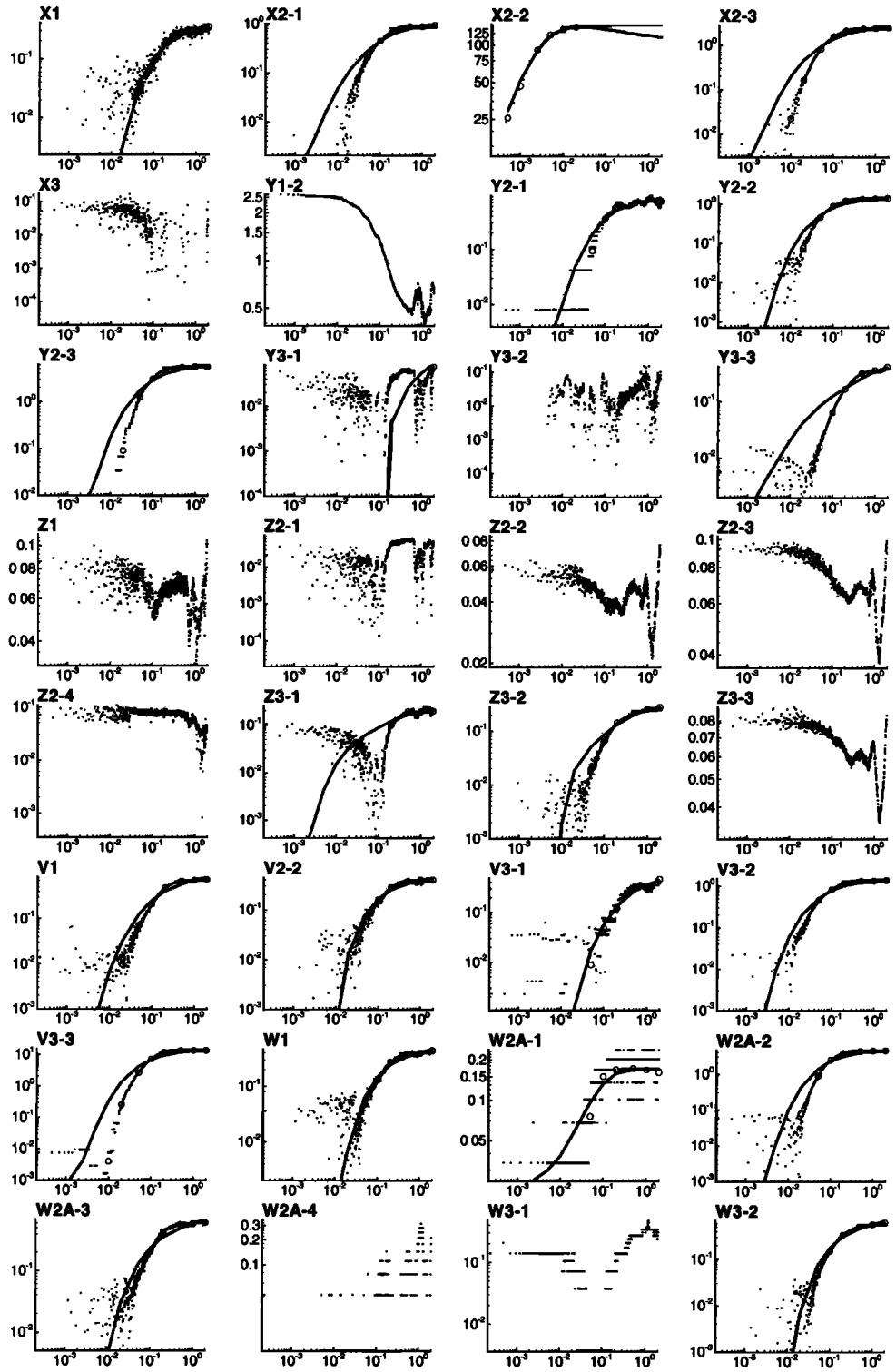


**Figure 2.** Observed (small dots) and simulated (curves, uniform rock) pressure buildup (kPa) versus time (days) during test PP5. Open circles are match points.

100 kPa, equal to the average at sea level. *Illman and Neuman* [2001] used analytically derived type curves to interpret all individual pressure records (except those in X3 and Y2-2) from cross-hole test PP4 by adopting the same barometric pressure. This allows us to compare our results with theirs (Figure 4). Though their estimates of $k$ consistently exceed ours by a factor of about 1.7, the logarithms of these two sets of estimates correlate quite well if one ignores intervals Z1, Z2-2, Z2-3, Z2-4, and Z3-3 (which, as discussed earlier, are located in a highly permeable zone that communicates pneumatically with the atmosphere). The relatively large type curve estimates of $k$ stem in part from a need to compensate for the disregard of airflow through high-permeability borehole intervals in the analytical solution. Indeed, when we repeat our numerical in-

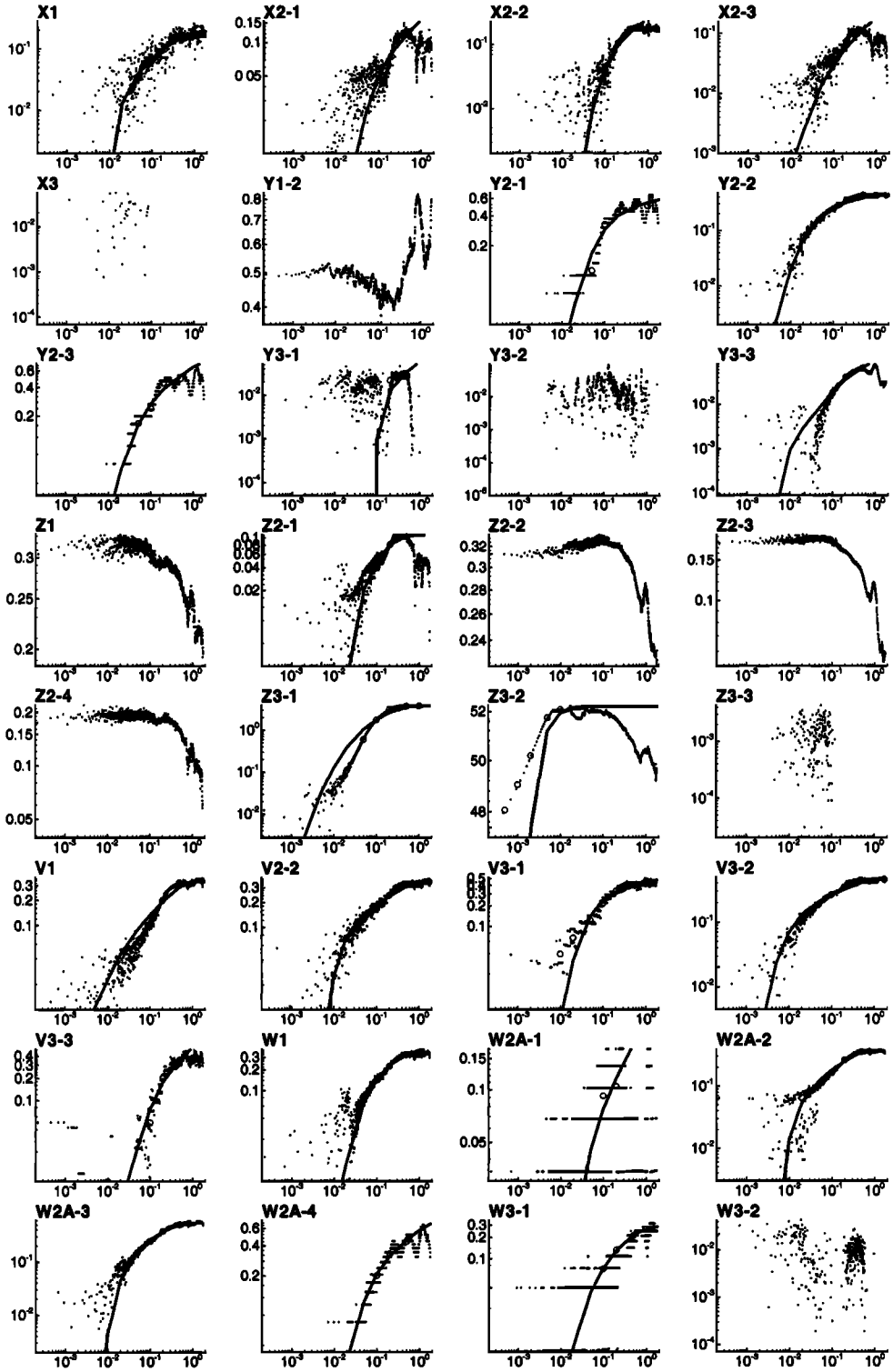


**Figure 3.** Observed (small dots) and simulated (curves, uniform rock) pressure buildup (kPa) versus time (days) during test PP6. Open circles are match points.

verse analysis without including the effect of packed-off borehole intervals, the ratio between the type-curve and numerical results drops down to 1.4. We show in Appendix A that it could be reduced further by accounting for air pressure within the injection interval in the linearized equations used by *Illman and Neuman* [2001].

There is good overall agreement between log-transformed estimates of $\phi$ obtained from type curves and estimates of $\log_{10}$ $\phi$ obtained by our numerical inverse method, except for intervals Y2-3, Y3-1, Z2-1, Z2-2, Z2-3, Z2-4, Z3-3, and V3-3 (Figure 4). Whereas the Z intervals are located in a highly permeable zone that communicates with the atmosphere, pressure transients in intervals Y2-3 and V3-3 are masked by borehole storage, which prevents the reliable estimation of porosity. The

**Table 1.** Parameters Identified for Cross-Hole Test PP4 by Treating the Medium as Spatially Uniform

| Interval | $\log_{10} k$, $m^2$ | $\log_{10} \phi$, $m^3/m^3$ |
|---|---|---|
| X1 | −14.08 ± 0.007 | −2.12 ± 0.021 |
| X2-1 | −13.94 ± 0.017 | −1.59 ± 0.054 |
| X2-2 | −14.26 ± 0.050 | −1.93 ± 0.170 |
| X2-3 | −14.06 ± 0.054 | −2.04 ± 0.131 |
| X3 | −12.18 ± 0.032 | −0.30 ± 0.111 |
| Y1-2 | −14.02 ± 0.017 | −1.21 ± 0.047 |
| Y2-1 | −14.03 ± 0.039 | −1.20 ± 0.140 |
| Y2-2 | −13.68 ± 0.001 | −1.22 ± 0.049 |
| Y2-3 | −14.30 ± 0.073 | −1.70 ± 0.266 |
| Y3-1 | −13.24 ± 0.015 | −1.64 ± 0.018 |
| Y3-2 | −13.18 ± 0.012 | −1.01 ± 0.047 |
| Y3-3 | −13.68 ± 0.032 | −1.19 ± 0.127 |
| Z1 | −12.42 ± 0.029 | −1.22 ± 0.028 |
| Z2-1 | −13.19 ± 0.047 | −1.28 ± 0.151 |
| Z2-2 | −12.62 ± 0.018 | −1.47 ± 0.027 |
| Z2-3 | −11.52 ± 0.090 | −0.30 ± 0.080 |
| Z2-4 | −11.92 ± 1.528 | −0.30 ± 0.670 |
| Z3-1 | −13.87 ± 0.017 | −0.99 ± 0.035 |
| Z3-2 | −13.80 ± 0.015 | −1.39 ± 0.037 |
| Z3-3 | −12.58 ± 0.099 | −0.52 ± 0.101 |
| V1 | −14.11 ± 0.027 | −1.90 ± 0.150 |
| V2-2 | −13.95 ± 0.021 | −1.94 ± 0.075 |
| V3-1 | −13.65 ± 0.014 | −1.45 ± 0.058 |
| V3-2 | −14.21 ± 0.007 | −2.29 ± 0.066 |
| V3-3 | −14.42 ± 0.041 | −1.62 ± 0.109 |
| W1 | −13.94 ± 0.025 | −1.82 ± 0.077 |
| W2A-1 | −13.36 ± 0.011 | −1.55 ± 0.042 |
| W2A-2 | −14.11 ± 0.041 | −1.48 ± 0.119 |
| W2A-3 | −13.82 ± 0.018 | −1.35 ± 0.043 |
| W2A-4 | −13.93 ± 0.024 | −1.22 ± 0.050 |
| W3-1 | −14.21 ± 0.025 | −2.08 ± 0.038 |
| W3-2 | −14.08 ± 0.048 | −2.08 ± 0.114 |

good agreement is due to the fact that air-filled porosity estimates are not sensitive to barometric pressure, as is pointed out in the following paragraph and explained in Appendix A.

In reality, the site is located at an elevation of 1200 m above sea level at which the average barometric pressure is not 100 kPa but is closer to 87 kPa. Repeating our analysis of pressure data from test PP4 with this latter value produces estimates of $k$ that are larger by a factor of ~1.1 but has virtually no effect on estimates of $\phi$ (except in a few cases associated with poor matches). The two sets of estimates are compared in Figure 5. *Illman and Neuman* [2001] concluded that a decrease in barometric pressure has no effect on estimates of $k$ but causes the estimates of $\phi$ to decrease by a factor of 1.1. We show in Appendix A that their conclusion, which is contrary to ours, stems from an artifact of their linearized analysis that is easy to eliminate. Appendix A explains why the factor of 1.1 is close to the ratio 100/87 = 1.16 between the two barometric pressures. It shows how estimates of $\log_{10} k$ can be modified a posteriori to take into account various values of average barometric pressure at the surface.

Attempts to interpret simultaneously pressure records from multiple test intervals, while treating the rock as being uniform, have led to unacceptably poor matches [*Illman et al.*, 1998]. The only way to improve these simultaneous matches is to account for rock heterogeneity in the numerical inverse model at a relatively high level of spatial resolution. This is what we proceed to do in section 4.

## 4. High-Resolution Inverse Estimation of Spatially Varying Parameters

Having obtained a statistical sample of equivalent pneumatic parameters for the heterogeneous fracture system at ALRS, we now turn to estimating the spatial variability of $\log_{10} k$ and $\log_{10} \phi$ at the site based on pressure data from cross-hole tests PP4, PP5, and PP6. We do so by analyzing simultaneously all pressure records either from one test at a time or from all three tests simultaneously. As explained in paper 1, we interpolate $\log_{10} k$ and $\log_{10} \phi$ by ordinary kriging (independently of each other) between estimates at a discrete set of "pilot points." Interpolation is done with the aid of 32 to 72 pilot points on a grid of cubic cells measuring 1 $m^3$. Most pilot points are placed along borehole monitoring intervals, but some are located between these and the injection interval. Geostatistical analysis has indicated [*Chen et al.*, 2000] that log permeabilities from single-hole tests can be associated with a directionally averaged power variogram, implying that they form a random fractal field of the fractional Brownian motion (fBm) type. On this basis we associate both log permeability and log porosity with power variograms having exponents $\beta$ that we estimate by our inverse method.

We found that estimating $\beta$ jointly with pneumatic parameter values at the pilot points would lead to simulation difficulties when $\beta$ exceeds 1. We avoided the problem by estimating sets of pneumatic parameters for a series of specified $\beta$ values between 0.5 and 1.75. Among these sets of estimates, we select those that provide the best trade-off between our objectives of minimizing the sum of squared residuals $\Phi$ and maintaining $\beta$ as close as possible to 0.45, the value previously obtained for log permeabilities from single-hole test data [*Chen et al.*, 2000]. Varying the two values of $\beta$ associated with $\log_{10} k$ and $\log_{10} \phi$ one at a time, by 0.25 about the optimum, produced changes in the pneumatic parameter estimates that were not large enough to justify searching for two separate optimum values of $\beta$.

**Table 2.** Summary Statistics of $\log_{10} k$ Estimated From Cross-Hole Tests by Treating the Rock as Being Spatially Uniform[a]

| Test | Sample Size | Minimum | Maximum | Mean | Variance | CV |
|---|---|---|---|---|---|---|
| PP4 | 32 | −14.42 | −11.52 | −13.57 | 0.568 | −0.0555 |
| PP5 | 21 | −15.15 | −12.22 | −13.94 | 0.397 | −0.0452 |
| PP6 | 23 | −14.4 | −12.0 | −13.66 | 0.220 | −0.0343 |
| PP7 | 29 | −15.61 | −12.0 | −13.67 | 0.560 | −0.0562 |
| PP8 | 31 | −14.57 | −11.5 | −13.68 | 0.530 | −0.0533 |
| All | 136 | −15.61 | −11.5 | −13.69 | 0.479 | −0.0505 |

[a]Units of $m^2$. CV, coefficient of variation.

**Table 3.** Summary Statistics of $\log_{10} \phi$ Estimated From Cross-Hole Tests by Treating the Rock as Being Spatially Uniform[a]

| Test | Sample Size | Minimum | Maximum | Mean | Variance | CV |
|---|---|---|---|---|---|---|
| PP4 | 29 | −2.29 | −0.52 | −1.53 | 0.168 | −0.267 |
| PP5 | 21 | −2.92 | −0.98 | −1.81 | 0.279 | −0.292 |
| PP6 | 22 | −2.64 | −0.58 | −1.80 | 0.237 | −0.271 |
| PP7 | 26 | −2.29 | −0.49 | −1.67 | 0.254 | −0.302 |
| PP8 | 30 | −2.31 | −0.33 | −1.46 | 0.225 | −0.326 |
| All | 128 | −2.92 | −0.33 | −1.63 | 0.241 | −0.301 |

[a]Units of $m^3/m^3$.

To initiate the iterative inverse process, we set $\log_{10} k$ and $\log_{10} \phi$ everywhere equal to the geometric mean of their previously estimated equivalent uniform values. Decreasing these initial values by an order of magnitude, or varying them randomly through the domain (based on uniform distributions ranging from −16 to −13 for $\log_{10} k$ and −3 to −1 for $\log_{10} \phi$) proved to have only a minor effect on the final estimates.

All numerical inversions converged in <50 iterations. The inversions required between 1000 and 5000 forward simulation runs that took 10 to 50 hours on the University of Arizona SGI Origin 2000 supercomputer with 32 processors.

### 4.1. Simultaneous Inversion of Data From Cross-Hole Test PP4

We start by discussing the simultaneous interpretation of pressures recorded in 32 borehole intervals during cross-hole test PP4, with one pilot point placed in each of these intervals (Figure 6). Most pilot points are placed at interval centers, but some are located off center. In guard intervals Y2-1 and Y2-3, they are offset toward the injection interval Y2-2 in order to enhance resolution near the latter. In the long interval V1 a pilot point is placed close to the ground surface where most of the airflow takes place.

Figure 7 shows how the normalized sum of squared residuals $\Phi$ varies with the variogram exponent $\beta$ in this case (dashed curve). Normalization is done with respect to the smallest $\Phi$ value in Figure 7. Normalized $\Phi$ is seen to decrease rapidly as $\beta$ increases to 1, then to remain relatively stable. This implies that the quality of model fit is relatively insensitive to $\beta$ when the latter exceeds 1. Considering this and our aim of achieving the best trade-off between minimizing the sum of squared residuals $\Phi$ and maintaining $\beta$ as close as possible to 0.45, we adopt the estimates corresponding to $\beta = 1$ as our optimum choice.

Figure 8 compares simulated and recorded pressure buildups versus time in all 32 intervals during cross-hole test PP4. Most of the simulated responses capture with reasonable fidelity the observed pressure behaviors in these intervals. Linear regression of measured versus simulated pressures yields a correlation coefficient $r^2$ equal to 0.998. The corresponding sum of squared residuals, $\Phi$, is 131.2 $kPa^2$, which is close to the value of 104.4 $kPa^2$ obtained earlier during the matching of individual pressure records to simulations conducted by treating the rock as being uniform (Figure 1). However, the simultaneous matches in intervals X2-2, Y2-3, Y3-2, Z3-2, V3-3, and W1 are noticeably better than were the corresponding individual matches. Whereas the simultaneous matches underestimate or overestimate observed pressure responses in intervals Y3-1, Z1, Z2-1, Z2-2, Z2-3, Z2-4, and Z3-3, they nevertheless reproduce the time rates of these responses more accurately than did the individual matches.

Estimates of $\log_{10} k$ and $\log_{10} \phi$ at pilot points, as well as associated individual 95% confidence limits, are listed in Table 4. In Table 4, pilot points are identified by the intervals in which they are located. Normalized log sensitivity to $k$ and $\phi$ at a given pilot point is computed by adding all terms in the corresponding column of a sensitivity matrix that corresponds to the optimal estimates, dividing by the largest sum among all pilot points (which corresponds to the estimate of $\log_{10} k$ at injection interval Y2-2), and taking logarithm to base 10. The results for both parameters are illustrated in Figure 9. The

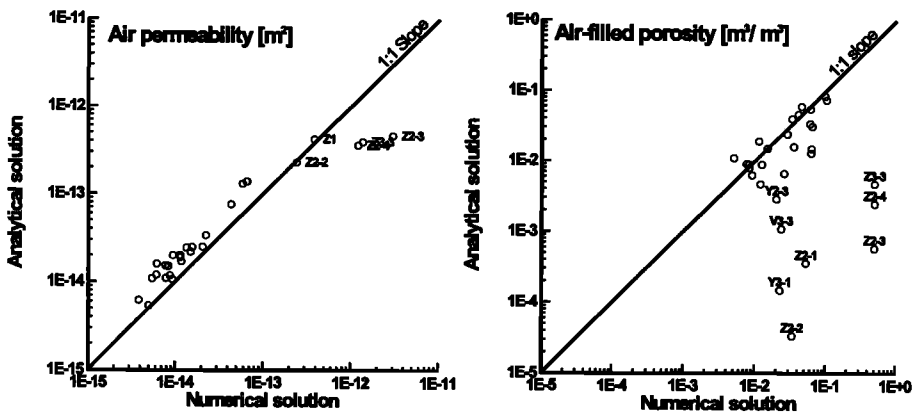


**Figure 4.** Pneumatic properties estimated analytically and numerically from PP4 test data.

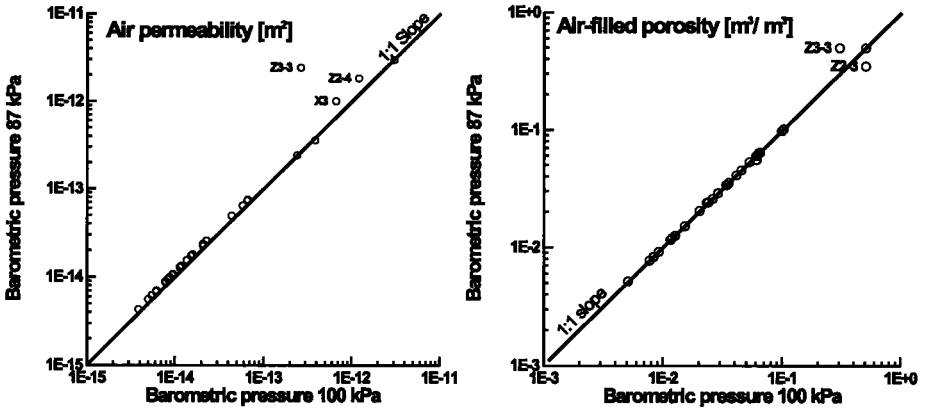


**Figure 5.** Estimates of pneumatic properties obtained for two reference barometric pressures from test PP4.

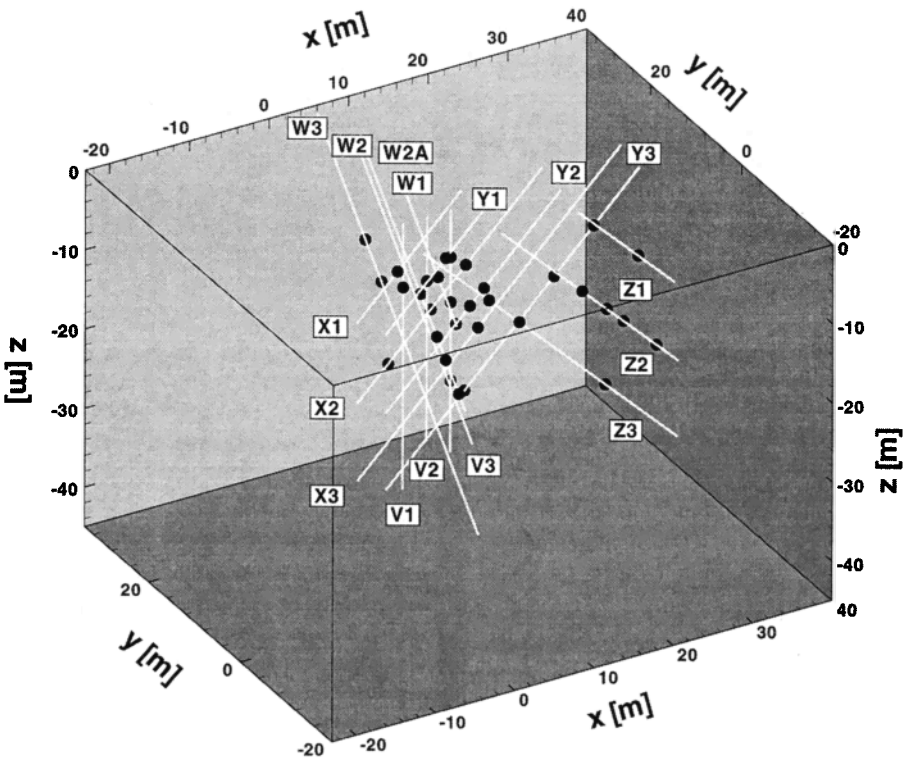


**Figure 6.** Three-dimensional representation of 32 pilot points in simulation domain.

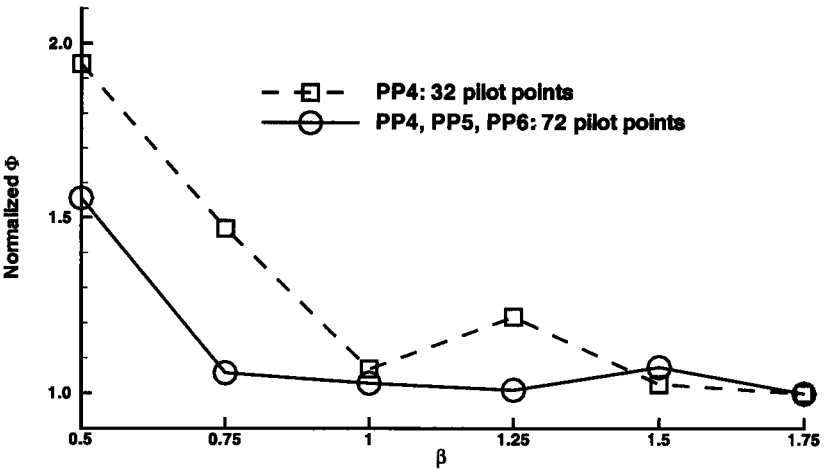


**Figure 7.** Sums of squared residuals versus exponent of power variogram, normalized by smallest value.

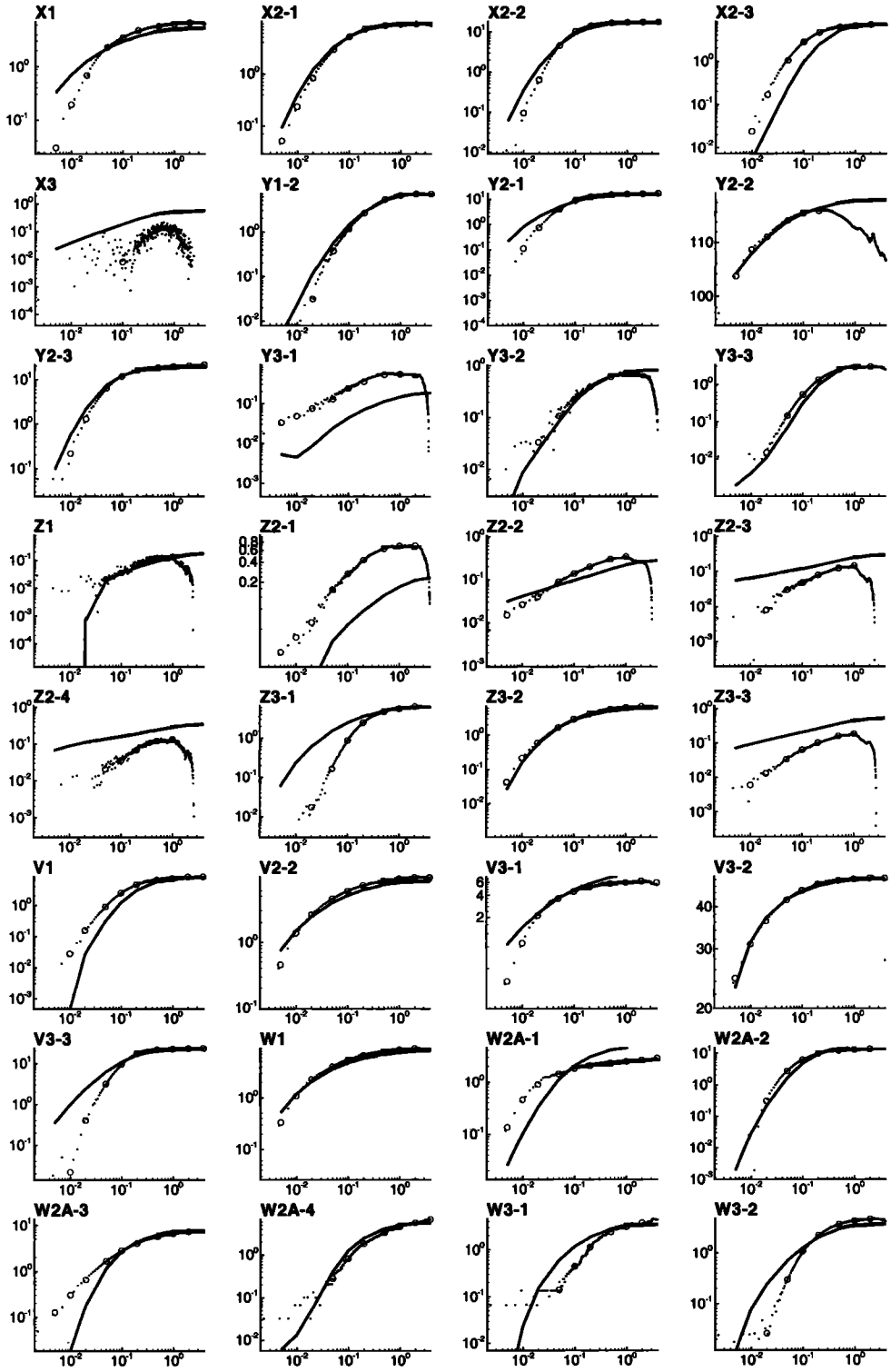


**Figure 8.** Observed (small dots) and simulated (curves, nonuniform rock with 32 pilot points) pressure buildup (kPa) versus time (days) during test PP4. Open circles are match points.

normalized log sensitivities of $\log_{10} k$ are seen to be larger than those of $\log_{10} \phi$, suggesting that estimates of $\log_{10} k$ are more reliable than those of $\log_{10} \phi$.

In Table 4, three estimates of $\log_{10} k$ coincide with the specified lower limit of $-20$ and one with the specified upper limit of $-10$. Nine estimates of $\log_{10} \phi$ coincide with the specified lower limit of $-5.0$ and three with the specified upper limit of $-0.3$. Some of the latter $\log_{10} \phi$ estimates are associated with low normalized log sensitivities (Figure 9) and relatively large confidence intervals (Table 4), suggesting that they are uncertain. Some are associated with high normalized log sensitivities (Figure 9), a fact we attribute to suboptimal pa-

**Table 4.** Parameters Identified at the Pilot Points by Numerical Inversion of Data From Cross-Hole Test PP4 Using 32 Pilot Points

| Interval | $\log_{10} k$, $m^2$ | $\log_{10} \phi$, $m^3/m^3$ |
|---|---|---|
| X1 | −11.60 ± 0.23 | −2.17 ± 0.87 |
| X2-1 | −15.27 ± 0.17 | −2.40 ± 0.22 |
| X2-2 | −15.29 ± 0.27 | −3.81 ± 0.48 |
| X2-3 | −16.01 ± 0.17 | −5.00 ± 0.45 |
| X3 | −20.00 ± 0.30 | −5.00 ± 0.97 |
| Y1 | −17.76 ± 0.32 | −2.72 ± 0.33 |
| Y2-1 | −16.61 ± 0.17 | −2.53 ± 0.40 |
| Y2-2 | −13.05 ± 0.03 | −2.38 ± 0.27 |
| Y2-3 | −14.56 ± 0.07 | −0.30 ± 0.55 |
| Y3-1 | −12.95 ± 0.17 | −5.00 ± 0.70 |
| Y3-2 | −20.00 ± 0.25 | −5.00 ± 0.40 |
| Y3-3 | −20.00 ± 0.64 | −5.00 ± 1.16 |
| Z1 | −10.00 ± 0.37 | −1.25 ± 1.07 |
| Z2-1 | −19.06 ± 0.41 | −5.00 ± 0.55 |
| Z2-2 | −14.37 ± 0.33 | −1.63 ± 0.96 |
| Z2-3 | −14.42 ± 0.38 | −2.76 ± 0.79 |
| Z2-4 | −14.35 ± 0.44 | −1.84 ± 0.93 |
| Z3-1 | −12.40 ± 0.41 | −0.30 ± 0.46 |
| Z3-2 | −11.93 ± 0.20 | −0.41 ± 0.25 |
| Z3-3 | −13.15 ± 0.36 | −0.30 ± 1.15 |
| V1 | −18.62 ± 0.23 | −5.00 ± 0.78 |
| V2 | −15.07 ± 0.37 | −4.59 ± 0.94 |
| V3-1 | −11.29 ± 0.18 | −1.02 ± 0.39 |
| V3-2 | −15.44 ± 0.07 | −4.23 ± 0.16 |
| V3-3 | −14.45 ± 0.66 | −1.17 ± 1.15 |
| W1 | −12.96 ± 0.74 | −1.68 ± 1.43 |
| W2-1 | −18.93 ± 0.46 | −5.00 ± 0.76 |
| W2-2 | −16.70 ± 0.15 | −4.18 ± 0.51 |
| W2-3 | −14.18 ± 0.18 | −2.18 ± 0.37 |
| W2-4 | −13.81 ± 0.14 | −1.65 ± 0.44 |
| W3-1 | −16.61 ± 0.03 | −2.88 ± 0.58 |
| W3-2 | −14.60 ± 0.19 | −5.00 ± 0.51 |

rameterization (number and spatial location of pilot points) and data (quality, number, and location of monitoring intervals).

Plate 1 depicts eigenvectors of the covariance matrix $\Sigma_a$ of parameter estimation errors associated with the optimum parameter estimates (as defined in paper 1). The eigenvectors are numbered so that the magnitudes of their corresponding eigenvalues increases with ascending numbers: the eigenvector with the smallest eigenvalue ($2 \times 10^{-7}$) is numbered 1, and that with the highest eigenvalue (1.3) is numbered 64. The first 32 components of each eigenvector correspond to $\log_{10} k$, and the next 32 components correspond to $\log_{10} \phi$, the two sets being separated by a horizontal line in Plate 1. The components of each eigenvector represent the relative contribution to it by the various parameter estimates. Parameters associated with eigenvectors that have small eigenvalues are less uncertain than those associated with eigenvectors that have large eigenvalues. Parameters associated with a single eigenvector have uncorrelated estimation errors. Parameters associated with multiple eigenvectors have cross-correlated estimation errors.

The first eigenvector in Plate 1 (corresponding to the smallest eigenvalue) is associated almost entirely with the eighth parameter, which is log permeability of the pilot point in the injection interval Y2-2. The corresponding estimate is therefore the least uncertain; indeed, it exhibits the largest normalized sensitivity (Figure 9) and the smallest confidence interval (Table 4). The last eigenvector (with the largest eigenvalue) is associated with a number of porosities that are cross-correlated and poorly estimated. Plate 1 demonstrates that eigenvectors associated with the smaller eigenvalues are dom-

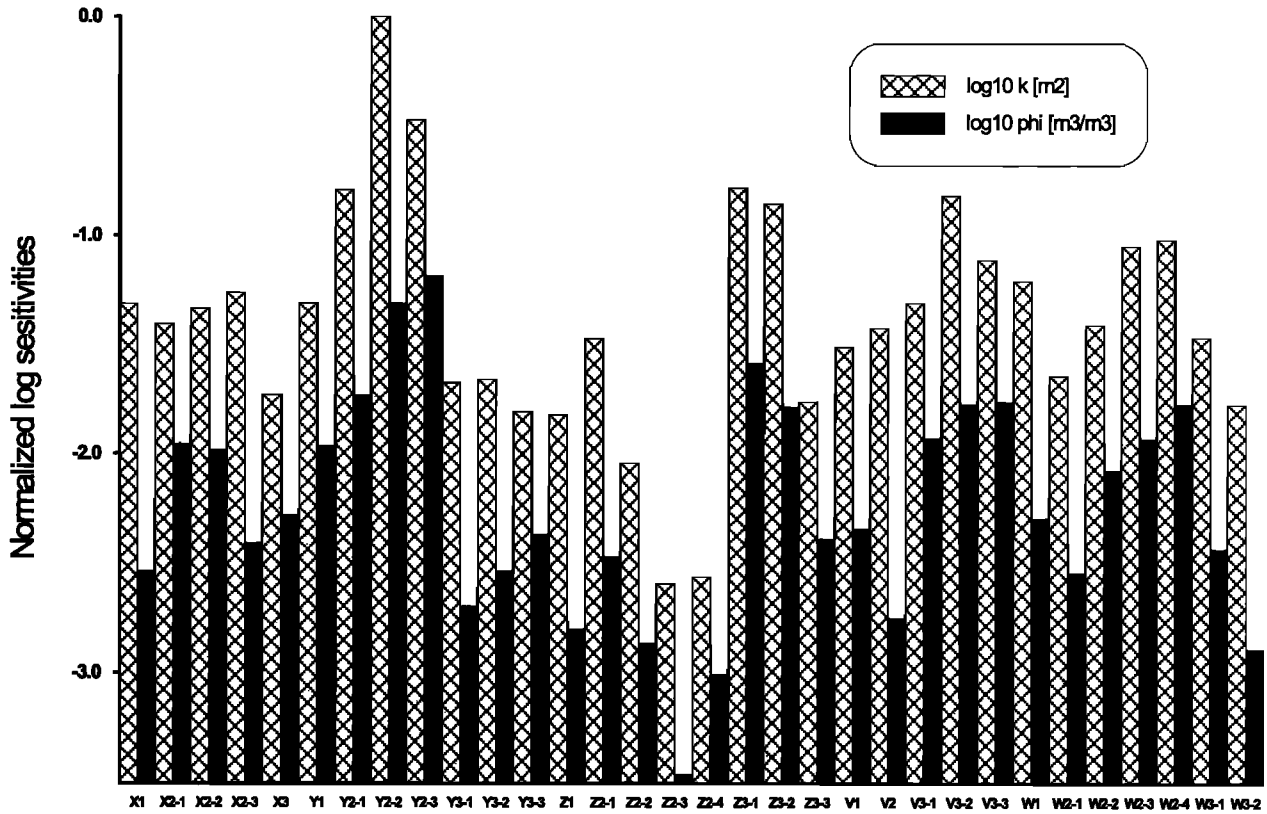


**Figure 9.** Normalized log sensitivities associated with parameter estimates for each pilot point obtained by simultaneous inversion of PP4 data.

inated by permeability estimates, while eigenvectors associated with the larger eigenvalues are dominated by porosity estimates. This indicates that log permeabilities are less uncertain than are log porosities.

Cross correlations between $\log_{10} k$ estimates, and cross correlations between $\log_{10} \phi$ estimates, are seen in Plate 1 to be generally much larger than those between $\log_{10} k$ and $\log_{10} \phi$ estimates. In other words, cross correlation between estimates of $\log_{10} k$ and $\log_{10} \phi$ is generally smaller than that between pairs of $\log_{10} k$ estimates and pairs of $\log_{10} \phi$ estimates. Some such cross correlation should be expected considering that pressure behavior depends in part on pneumatic diffusivity, which includes the ratio between permeability and porosity.

A three-dimensional view of $\log_{10} k$ estimates, obtained by kriging the above 32 pilot estimates over a three-dimensional grid of 1-m scale cubes, is depicted in Plate 2. Plate 3 depicts a corresponding view of kriged $\log_{10} \phi$ values. The two parameters are seen to exhibit comparable patterns of spatial variability.

Repeating the inversion with 64 pilot points, by adding 32 such points between and along test intervals [*Vesselinov*, 2000], improves the matches but increases estimation uncertainty (this is discussed further in section 4.3). Most importantly, it does not change in any significant way the spatial distribution of kriged $\log_{10} k$ and $\log_{10} \phi$ estimates.

We also analyzed simultaneously all pressure records from cross-hole test PP5 and cross-hole test PP6, using 32 pilot points in each case. The results yield tomographic images of $\log_{10} k$ and $\log_{10} \phi$ that are visually comparable to those obtained with 32 pilot points from test PP4 [*Vesselinov*, 2000]. This apparent similarity notwithstanding, using kriged estimates from test PP4 to simulate pressure responses during tests PP5, PP6, and PP7 results in numerous poor matches [*Vesselinov*, 2000]. This is true especially in pressure monitoring intervals situated close to the respective injection intervals. On the other hand, responses calculated in intervals close to borehole Y2 are quite satisfactory. This is so because during test PP4, air was injected into borehole Y2, and therefore rock properties near it have been defined with greater accuracy than elsewhere in the rock (Table 4 and Plate 1). To obtain a more even definition of spatial variability throughout the tested rock volume, we repeat the inversion by considering simultaneously three tests with injection intervals located in three different boreholes.

### 4.2. Simultaneous Inversion of Data From Cross-Hole Tests PP4, PP5, and PP6

During tests PP4, PP5, and PP6, air was injected at various rates into packed-off intervals in boreholes Y2, X2, and Z3, respectively. To invert simultaneously pressure data from all three tests, we found it useful to place 72 pilot points along monitoring intervals as well as between them and the injection intervals. The locations of these pilot points are shown in Figure 10.

Figure 7 shows how the sum of squared residuals $\Phi$, normalized with respect to the smallest $\Phi$ (corresponding to $\beta = 1.25$), varies with the power variogram exponent $\beta$ in this case (solid curve). The normalized $\Phi$ is seen to decrease rapidly as $\beta$ increases to 0.75, then to remain relatively stable. Though we consider the estimates corresponding to $\beta = 0.75$ to provide an optimum trade-off between our aims of minimizing $\Phi$ and $\beta$, we nevertheless consider below estimates that correspond to $\beta = 1$. We do so to allow their direct comparison with previous estimates obtained on the basis of data from test PP4 for $\beta = 1$. These estimates are very similar to those corresponding to the smallest value of $\Phi$ at $\beta = 1.25$ and to those corresponding to $\beta = 0.75$.

Figures 11, 12, and 13 compare simulated pressure buildups versus time with those recorded in various intervals during tests PP4, PP5, and PP6, respectively. We have ascribed match points to 76 of the 96 pressure records (32 per test) depicted in Figures 11–13. This notwithstanding, only a small number among the 96 matches are poor, while the majority are of intermediate, good to excellent quality. The sum of squared pressure residuals, $\Phi$, is 133.2 kPa$^2$ of which 45.4 is due to test PP4 (much less than the 131.2 value obtained earlier upon analyzing test PP4 alone with 32 pilot points), 82.7 due to test PP5, and 5.1 due to PP6. Matching individual pressure records while treating the rock as being uniform yields a $\Phi$ value equal to 173.9 kPa$^2$ of which 104.4 is contributed by test PP4, 68.8 by PP5, and 0.7 by PP6 [*Vesselinov*, 2000].

As in the case where we analyze simultaneously pressure data from test PP4 alone, most estimates of $\log_{10} k$ at pilot points are characterized by relatively large normalized log sensitivities, while those of $\log_{10} \phi$ have lower sensitivities [*Vesselinov*, 2000]. Both sets of estimates appear to be reasonable. They lead to kriged estimates of $\log_{10} k$ (Plate 4) and $\log_{10} \phi$ (Plate 5) that are similar in their overall appearance to those obtained previously through inversion of PP4 test data alone (Plates 2 and 3, respectively).

### 4.3. Discrimination Between Alternative Inversion Schemes

We discussed in some detail two sets of stochastic inverse results corresponding to data from test PP4 with 32 pilot points and data from tests PP4, PP5, and PP6 with 72 pilot points. We also briefly mentioned results corresponding to data from test PP4 with 64 pilot points and data from each of tests PP5 and PP6 with 32 pilot points. Other inversion schemes are possible, and it is therefore of interest to compare and rank them on the basis of some formal criteria.

In paper 1 we discussed four model discrimination criteria (AIC, BIC, $\phi_M$, and $d_M$) that we apply below to three of the aforementioned inversion schemes (test PP4 with 32 and 64 pilot points; tests PP4, PP5, and PP6 with 72 pilot points). We list these criteria in Table 5 together with other relevant statistics (defined in paper 1) including degrees of freedom (number of match points minus number of unknown parameters), $\Phi$ (sum of squared residual pressures), $s_p^2$ (posterior estimate of pressure measurement variance), $\log |\Sigma_a|$ (log-transformed determinant of estimation covariance matrix $\Sigma_a$), and $S$ (optimum log likelihood function). Rankings of the three models based on each of these criteria and statistics are shown in parentheses.

All four discrimination criteria identify the inversion scheme based on test PP4 with 64 pilot points as being the best and the scheme based on simultaneous interpretation of tests PP4, PP5, and PP6 with 72 pilot points as being the worst among the three schemes considered. On the sole basis of degrees of freedom, one would rank the latter scheme as being most promising. It, however, is associated with relatively large values of $\Phi$ and $S$, indicating that it leads to relatively poor matches with observed pressure data. We attribute these poor matches, and the associated low ranking, of the scheme to the large signal-to-noise ratio associated with test PP4 as compared to tests PP5 and PP6. In other words, pressure data from tests PP5 and PP6 are noisy in comparison to those from test PP4

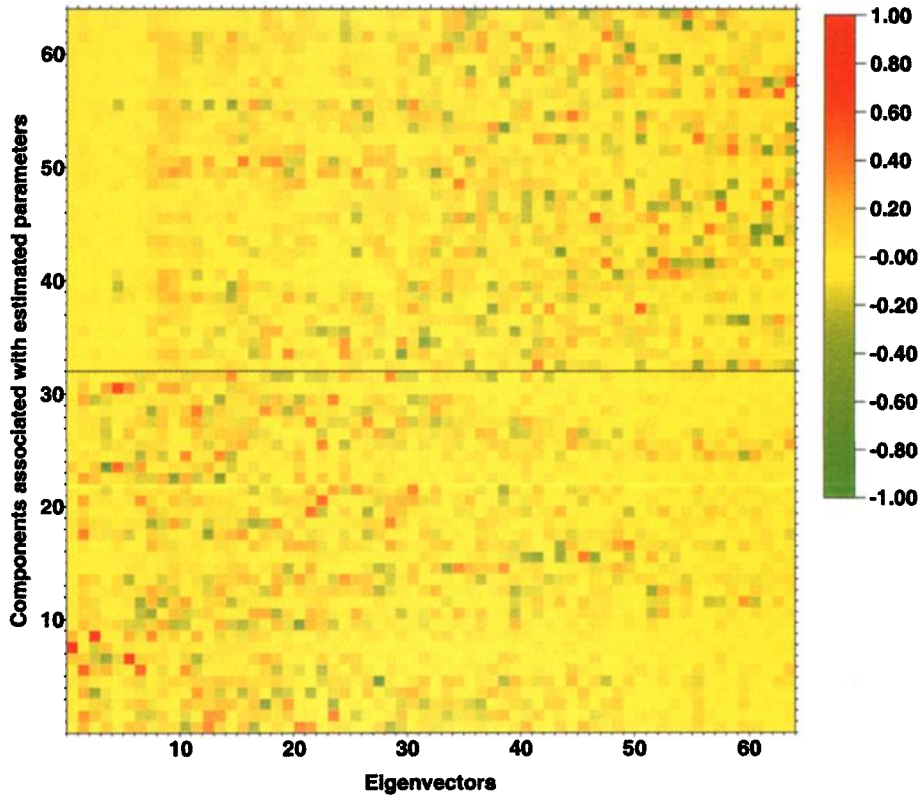


**Plate 1.** Two-dimensional representation of eigenvectors of normal matrix of estimation errors at pilot points obtained by simultaneous inversion of PP4 test data.

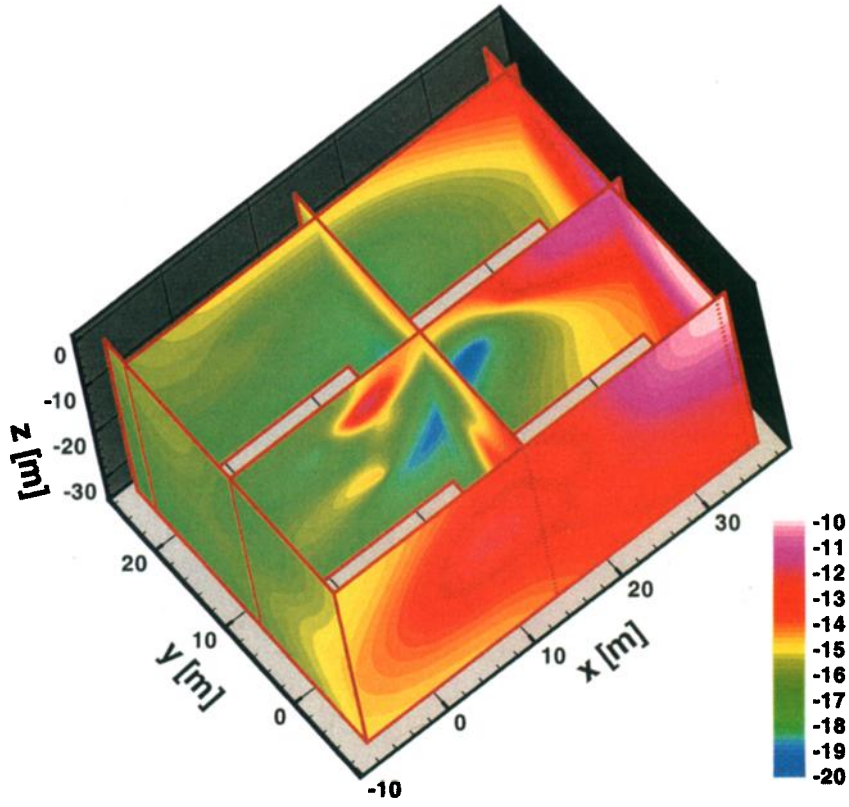


**Plate 2.** Three-dimensional representation of kriged $\log_{10} k$ [$m^2$] estimated by simultaneous inversion of PP4 test data with 32 pilot points.

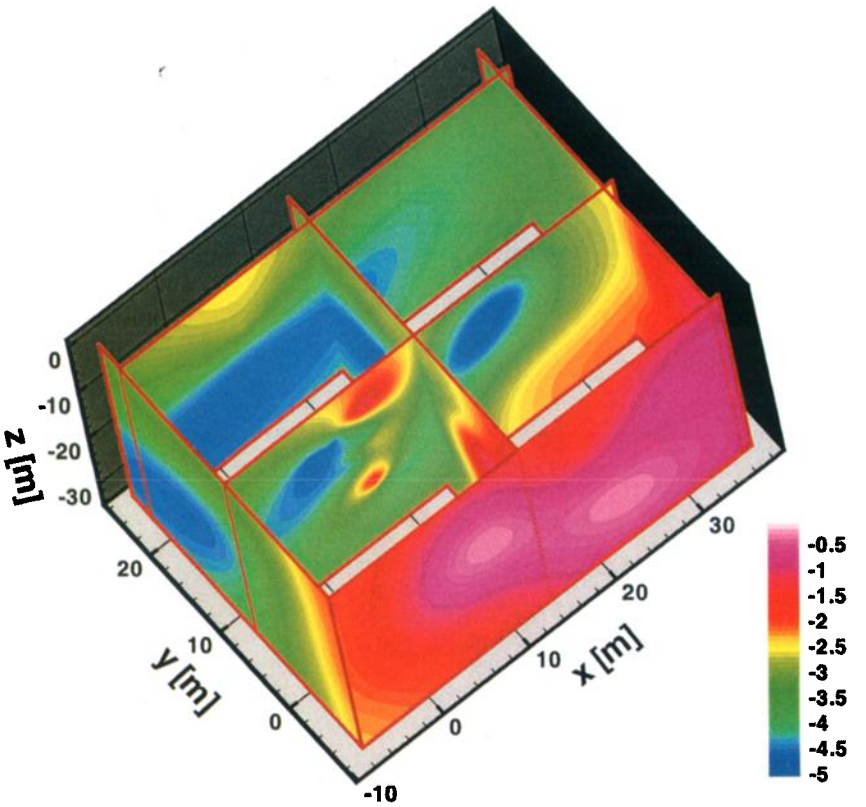


**Plate 3.** Three-dimensional representation of kriged $\log_{10} \phi$ estimated by simultaneous inversion of PP4 test data with 32 pilot points.

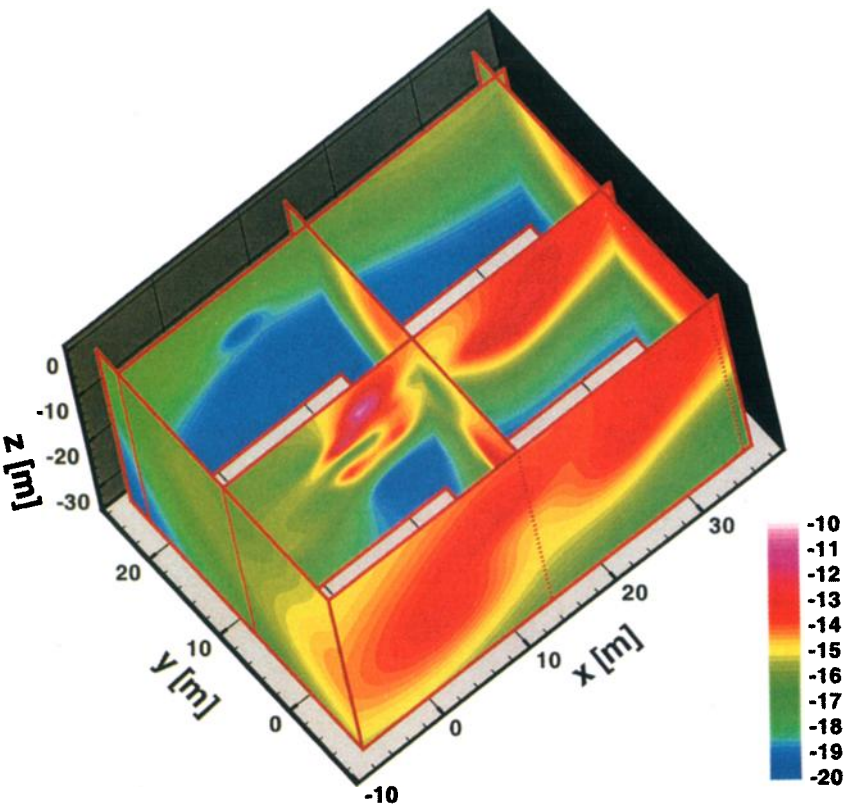


**Plate 4.** Three-dimensional representation of kriged $\log_{10} k$ [$m^2$] estimated by simultaneous inversion of PP4, PP5, and PP6 test data.

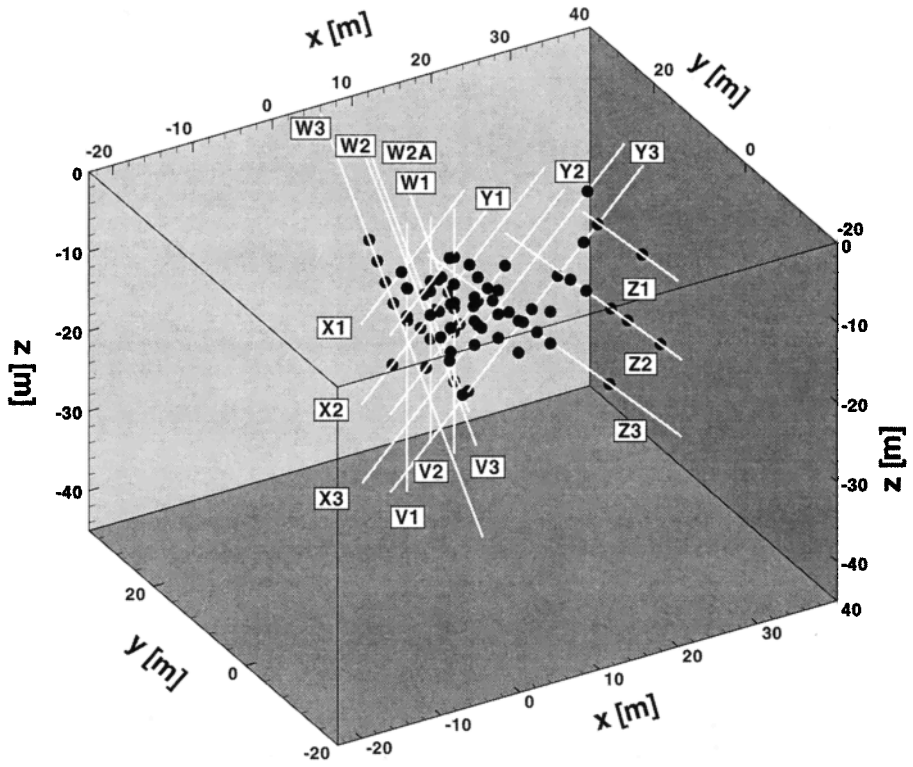


**Figure 10.** Three-dimensional representation of 72 pilot points within simulation domain.

(as is evident in Figures 2, 3, and 4 of paper 1) and therefore more difficult to reproduce with our model.

The reasons why test PP4 is more reliable than other cross-hole tests conducted at ALRS have been discussed in some detail by *Illman and Neuman* [2001]. We note in particular that during test PP4, injection took place into an extensive high-permeability zone at a relatively high injection rate, which produced well-defined pressure responses in most monitoring intervals. During test PP5, injection took place at a relatively low rate into a borehole interval embedded within a low-permeability zone, resulting in relatively weak pressure signals in monitoring intervals. During test PP6, injection took place at a relatively low rate into a borehole interval embedded within a high-permeability zone that appears to be connected to the atmosphere, resulting in attenuated and noisy pressure signals in monitoring intervals. This made it difficult to clearly separate signal from background noise during these two tests, especially in monitoring intervals located far from injection boreholes. We find it encouraging that all four model discrimination criteria in Table 5 recognize this by ranking the scheme that uses less reliable data from tests PP5 and PP6 as inferior to the two schemes that rely solely on superior data from test PP4.

The above findings must be tempered by our recognition that it would have been logical to assign lower weights to pressure data from tests PP5 and PP6 than to those from test PP4. Instead, we assigned equal weights to all the pressure match point values. We did so in the absence of clear statistical indicators to guide us in the assignment of variable weights. It follows that the statistics in Table 5 may not accurately reflect the noted differences in reliability between the various pressure data sets.

Of the two schemes that utilize only PP4 data, the one with 64 pilot points produces a much better representation of observed pressures than does the one with 32 pilot points. Therefore it is ranked first even though it has fewer degrees of freedom due to its larger number of unknown parameters, which render it less parsimonious.

Notwithstanding these considerations, the scheme based on test PP4 with 64 pilot points is associated with a much larger value of log $|\Sigma_a|$ than are the two other schemes. In other words, even though the scheme leads to very good matches between observed and computed pressures, it results in parameters that are far less certain than are those obtained by the other two schemes. When parameters obtained from test PP4 with 32 and 64 pilot points had been used to simulate pressures recorded during tests PP5–PP7, they yielded matches with comparable sums of squared residuals. However, the first set resulted in a better qualitative reproduction of temporal pressure variations during tests PP6 and PP7 [*Vesselinov*, 2000]. This explains why we have devoted more space in this paper to the two inverse schemes that, though ranked second and third by formal model discrimination criteria in Table 5, nevertheless lead to the most reliable estimates of model parameters.

## 5. Relationship Between Air Permeability Estimates From Cross-Hole and Single-Hole Tests

The tomographic images of kriged $\log_{10} k$ estimates that we obtain through simultaneous inversion of pressure data from cross-hole test PP4 (Plate 2) and cross-hole tests PP4, PP5, and PP6 (Plate 4) show striking similarities with the image obtained from single-hole test data by ordinary kriging (Plate 6 [*Vesselinov*, 2000]). All three images represent values on a nominal scale of $1 \times 1 \times 1$ m$^3$. The cross-hole tomographic images

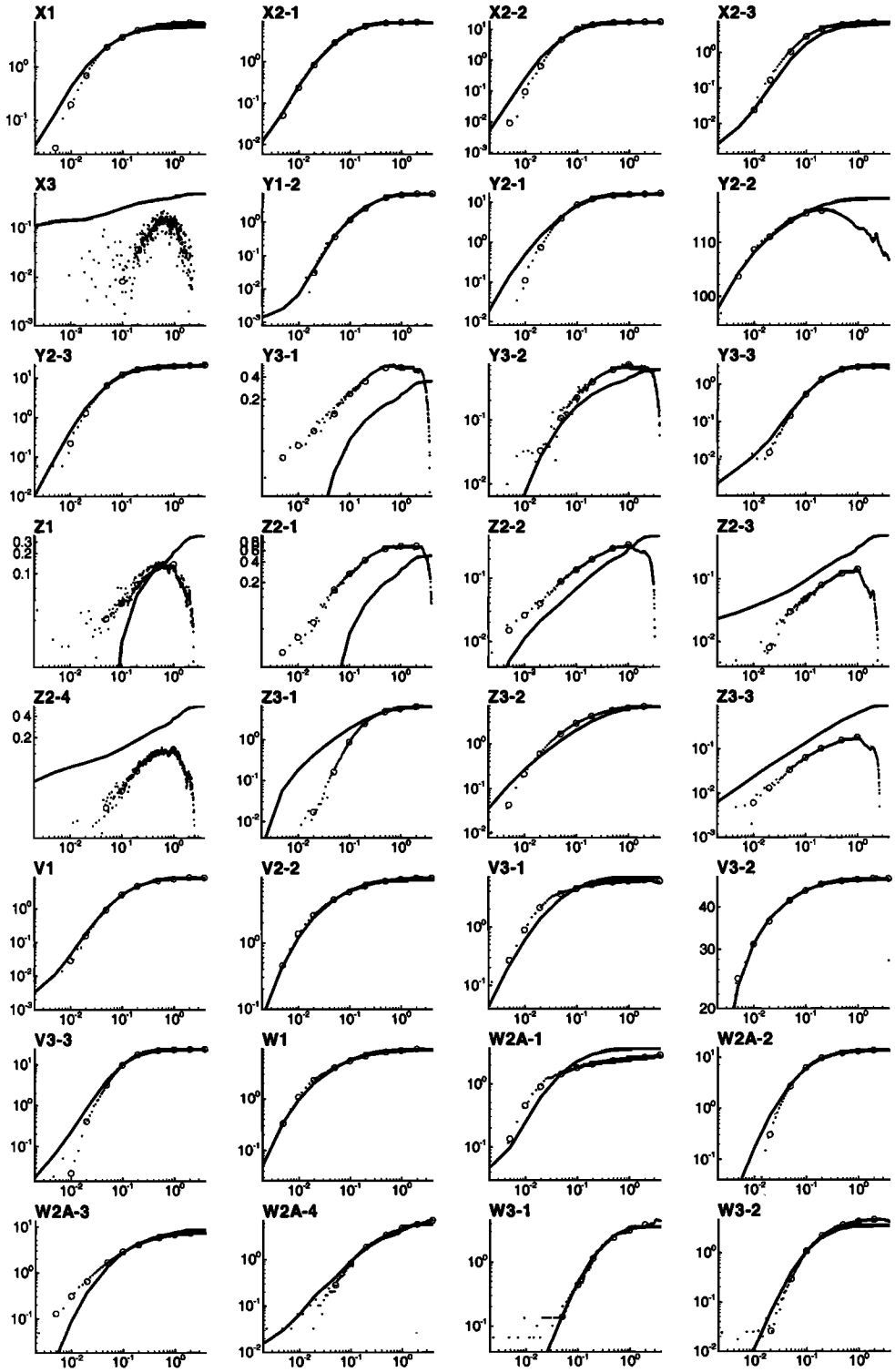


**Figure 11.** Observed (small dots) and simulated (curves, nonuniform rock with 72 pilot points) pressure buildup (kPa) versus time (days) for test PP4 obtained by simultaneous inversion of PP4, PP5 and PP6 test data.

show more detail, sharper contrasts, and a wider range of values than does the single-hole kriged image. We attribute this to the simultaneous analysis of numerous transient pressure records from cross-hole tests as compared to steady state analyses of individual single-hole tests and the relatively large rock volume affected by the cross-hole tests.

Figure 14 compares kriged estimates obtained via simultaneous inversion of data from cross-hole tests PP4, PP5, and PP6 with single-hole air permeabilities determined along four boreholes (X2, Y2, Y3, and V2). In borehole Y2 the two sets of $\log_{10} k$ values correlate quite closely even though there are many fewer pilot points along it (5) than single-hole test data

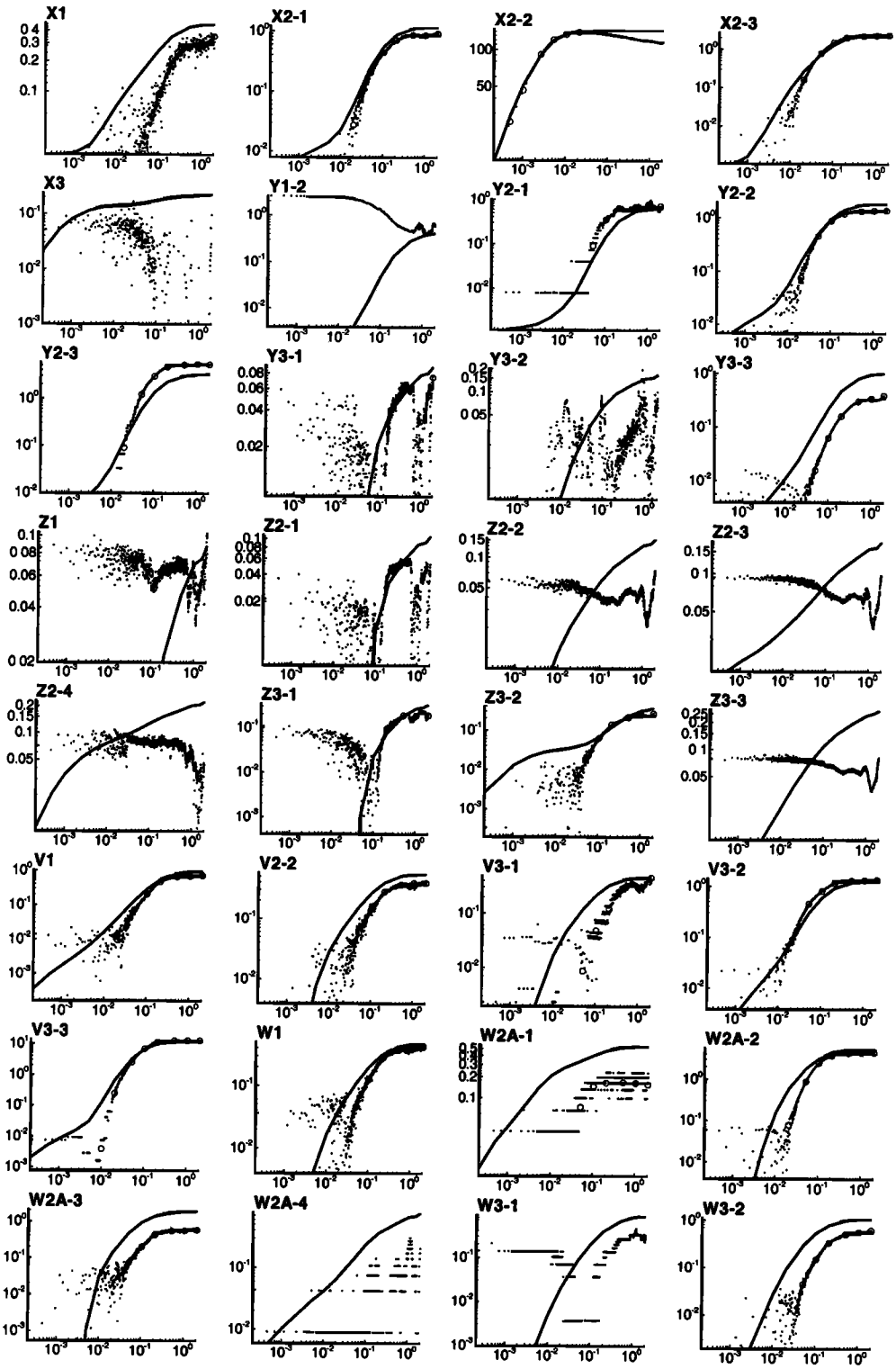


**Figure 12.** Observed (small dots) and simulated (curves, nonuniform rock with 72 pilot points) pressure buildup (kPa) versus time (days) for test PP5 obtained by simultaneous inversion of PP4, PP5 and PP6 test data.

(28). The same is true to a lesser extent for borehole V2, but the correlations are weaker in boreholes X2 and Y3. This may be due in part to the much smaller number of pilot points used in our inversion of the cross-hole test data than the number of measurements available from single-hole tests.

Table 6 lists the mean, variance, and coefficient of variation for these and some additional sets of kriged estimates at 53,176 nodes of our auxiliary grid. Kriged estimates from cross-hole tests have sample mean values ranging from −14.17 to −15.69 and averaging −15.07. These cross-hole mean estimates are

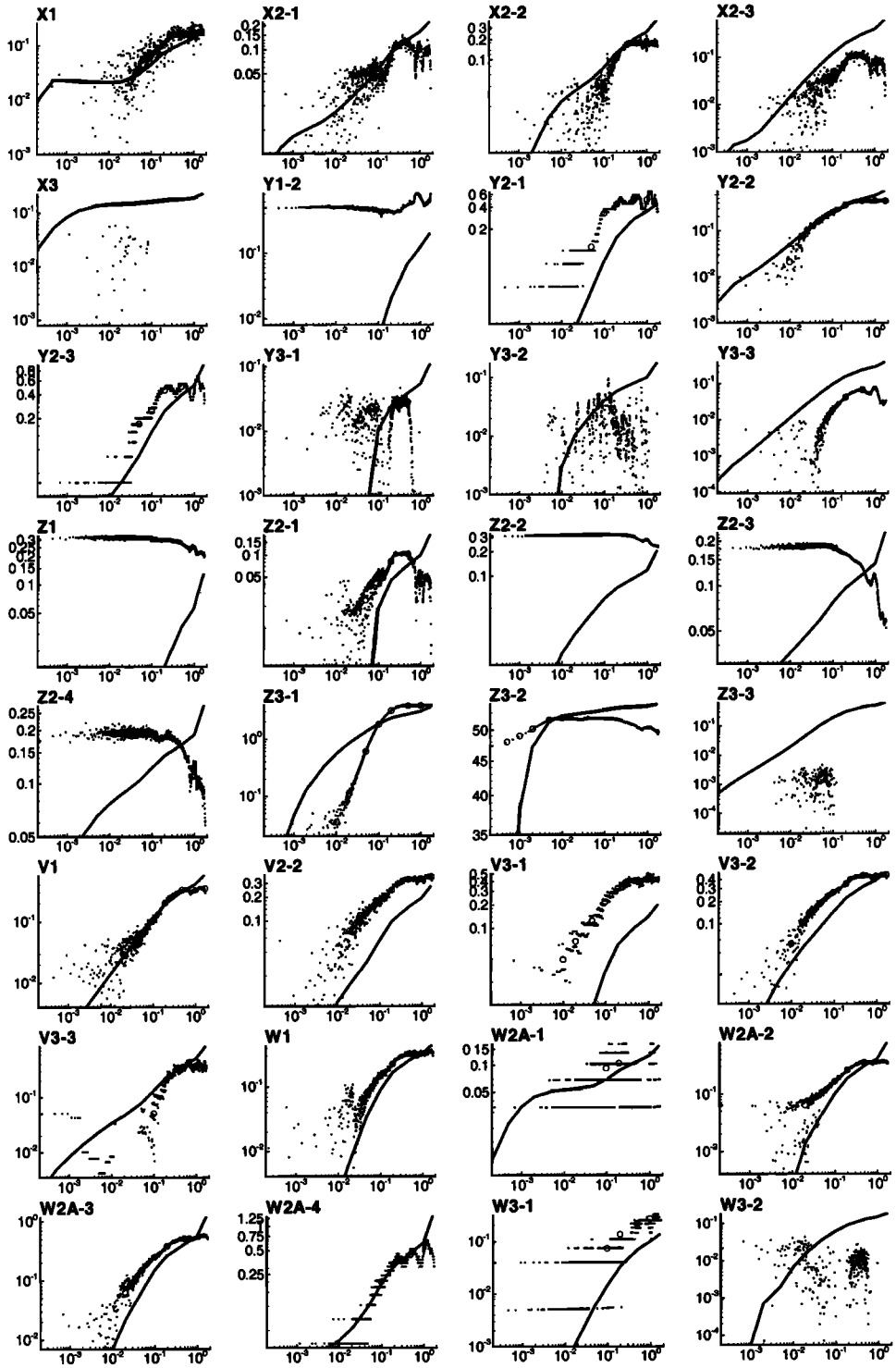


**Figure 13.** Observed (small dots) and simulated (curves, nonuniform rock with 72 pilot points) pressure buildup (kPa) versus time (days) for test PP6 obtained by simultaneous inversion of PP4, PP5 and PP6 test data.

close to each other and to the mean of single-hole kriged estimates, −15.20, indicating consistency in the mean between all kriged estimates from both types of tests. The sample variances of kriged sets obtained from cross-hole tests exceed those obtained from single-hole tests by factors ranging from 3.0 to 7.0 and averaging 4.7. We attribute this to estimation errors that stem from the inability of our inverse model to fully reproduce observed pressure behavior during cross-hole tests at ALRS, insufficient sensitivity of some of our estimates to the recorded pressure signals, and corruption of some of these signals by excessive noise in the data.

Table 6 also lists the mean, variance, and coefficient of variation of inverse $\log_{10} k$ estimates at pilot point locations based on cross-hole test data and values of $\log_{10} k$ corresponding to single-hole test data. Pilot point estimates from cross-hole tests have sample mean values ranging from

**Table 5.** Statistics, Model Discrimination Criteria, and Rankings of Three Inversion Schemes[a]

| Criteria | Models | | |
|---|---|---|---|
| | PP4 | PP4 | PP4, PP5, PP6 |
| Number of pilot points | 32 | 64 | 72 |
| Number of unknowns | 64 | 128 | 144 |
| Number of match points | 252 | 252 | 489 |
| Degrees of freedom | 188 (2) | 124 (3) | 345 (1) |
| $\Phi$ | 131.2 (2) | 17.03 (1) | 133.2 (3) |
| $s_p^2$ | 0.521 (3) | 0.068 (1) | 0.272 (2) |
| $\log \lvert\Sigma_a\rvert$ | −33 (1) | 3 (3) | −29 (2) |
| $S$ | 551 (2) | 36 (1) | 752 (3) |
| AIC | 679 (2) | 292 (1) | 1040 (3) |
| BIC | 905 (2) | 744 (1) | 1644 (3) |
| $\phi_M$ | 770 (2) | 474 (1) | 1277 (3) |
| $d_M$ | 820 (2) | 506 (1) | 1408 (3) |

[a]Rankings in parentheses.

−14.23 to −15.20 and averaging −14.97. These cross-hole mean estimates are close to each other and to the mean of single-hole test data, −15.22, indicating consistency in the mean between them. They are very close to the mean values of corresponding kriged estimates, as one should expect. The sample variances of pilot point estimates from cross-hole tests exceed those of the corresponding kriged estimates by factors ranging from 1.0 to 4.1 and averaging 2.5. The sample variance of the single-hole test data likewise exceeds that of the corresponding kriged estimates by a comparable factor of 1.9. This reflects the tendency of kriged estimates to vary smoothly between values one specifies at either pilot or measurement points. The sample variances of estimates obtained at pilot points from cross-hole tests exceed that of the single-hole test data by factors ranging from 1.6 to 8.2 and averaging 6.05. We again attribute this to estimation errors as explained in the preceding paragraph.

## 6. Relationship Between Estimates of Air-Filled Porosity and Air Permeability From Cross-Hole Tests

The kriged estimates of $\log_{10} \phi$ we obtain through simultaneous inversion of pressure data from cross-hole test PP4 (Plate 3) and cross-hole tests PP4, PP5, and PP6 (Plate 5) show similar spatial patterns. Both sets of estimates represent values on a nominal scale of $1 \times 1 \times 1$ m$^3$. We do not have at our disposal a sufficient number of porosity estimates from single-hole tests to compare with the former (four such estimates were obtained by *Vesselinov and Neuman* [2001] using a numerical inverse model). Table 7 allows one to compare the

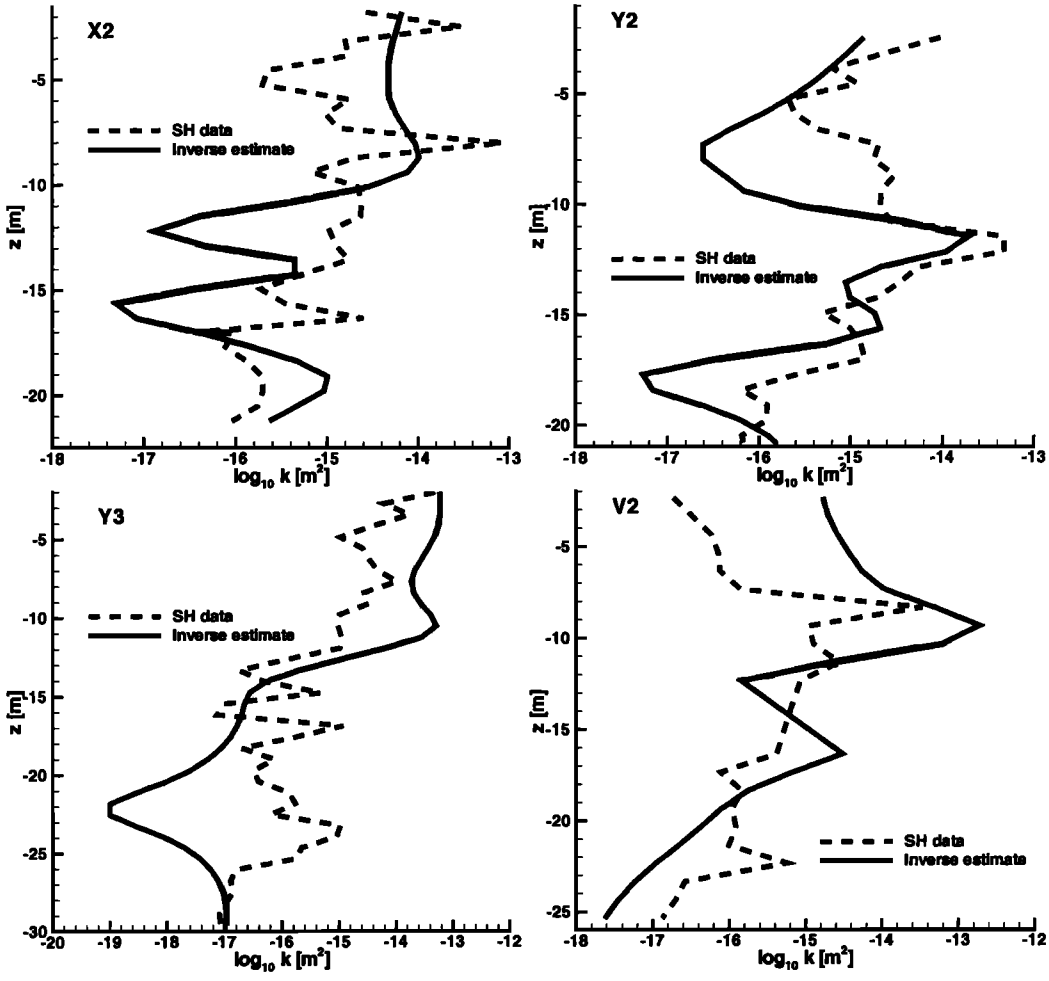


**Figure 14.** Air permeability along selected boreholes estimated from single-hole tests and by simultaneous inversion of PP4, PP5, and PP6 test data.

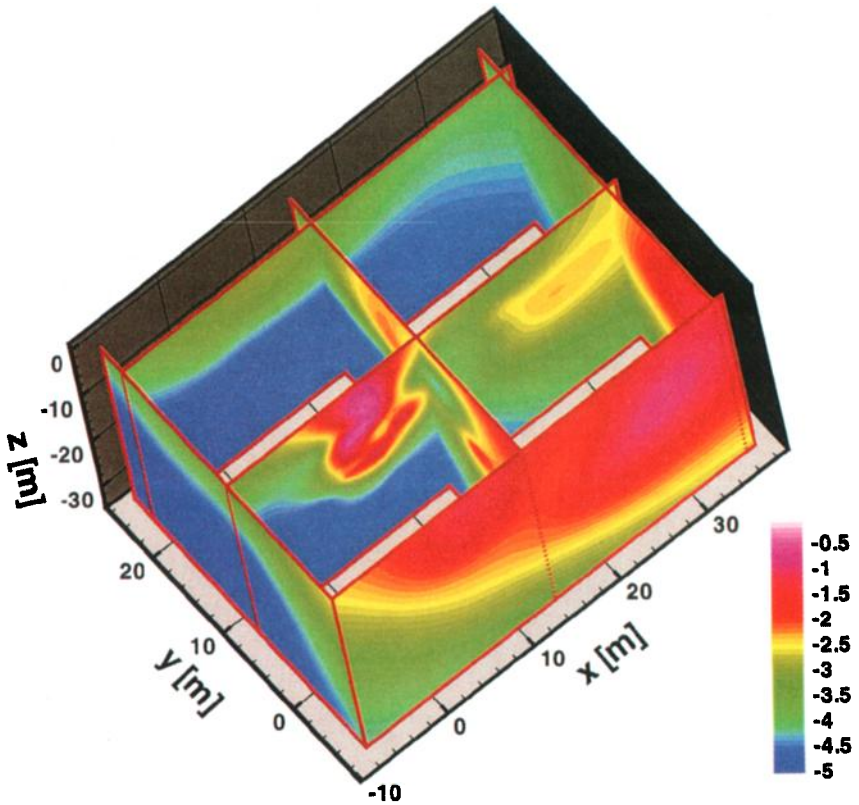


**Plate 5.** Three-dimensional representation of kriged $\log_{10} \phi$ estimated by simultaneous inversion of PP4, PP5, and PP6 test data.

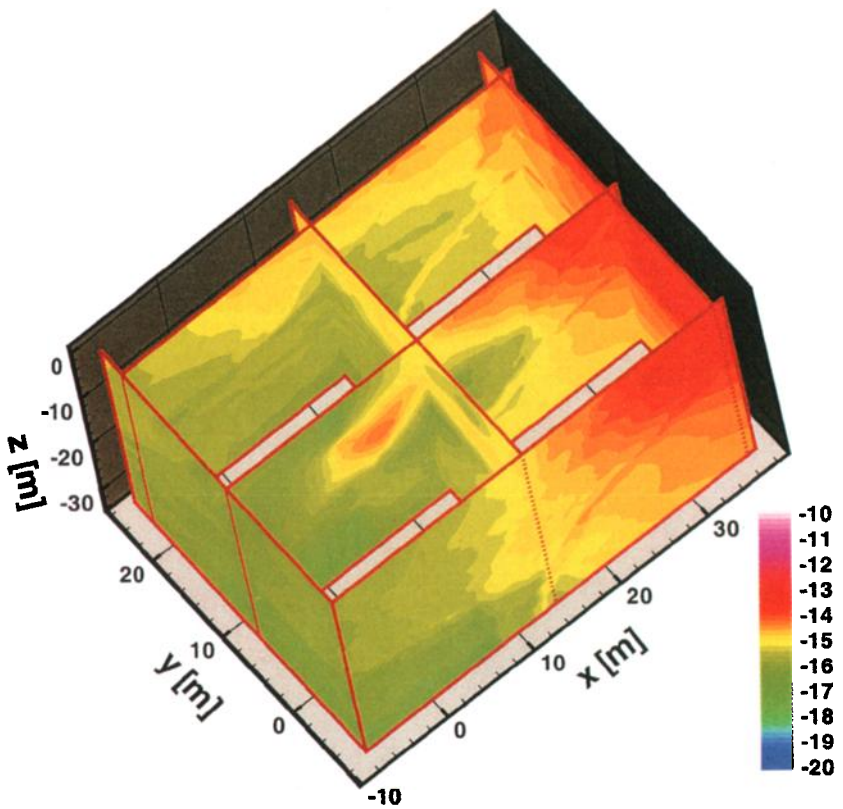


**Plate 6.** Three-dimensional representation of kriged $\log_{10} k$ [$m^2$] estimates based on single-hole test results.

**Table 6.** Summary Statistics of $\log_{10} k$ Estimated in Various Ways[a]

| Source | Data or Pilot Point Estimates | | | | Kriging Estimates | | | |
|---|---|---|---|---|---|---|---|---|
| | Sample Size | Mean | Variance | CV | Sample Size | Mean | Variance | CV |
| Single-hole tests (uniform rock) | 227 | −15.22 | 0.87 | −0.061 | 53176 | −15.20 | 0.45 | −0.044 |
| Cross-hole tests (type curve analyses) | | | | | | | | |
| PP4 [after *Illman and Neuman*, 2001] | 30 | −13.46 | 0.34 | −0.043 | | NA | | |
| Cross-hole tests (numerical analyses) | | | | | | | | |
| Uniform rock | | | | | | | | |
| PP4–PP8 | 136 | −13.69 | 0.48 | −0.051 | | NA | | |
| Nonuniform rock | | | | | | | | |
| PP4 | 32 | −15.17 | 7.17 | −0.176 | 53176 | −15.15 | 3.17 | −0.117 |
| PP4 | 64 | −15.07 | 4.72 | −0.144 | 53176 | −15.07 | 1.95 | −0.093 |
| PP5 | 32 | −15.19 | 6.79 | −0.171 | 53176 | −15.26 | 1.64 | −0.084 |
| PP6 | 32 | −14.23 | 1.37 | −0.082 | 53176 | −14.17 | 1.34 | −0.082 |
| PP4–PP6 | 72 | −15.20 | 6.25 | −0.164 | 53176 | −15.69 | 2.48 | −0.100 |

[a]Units of $m^2$. NA, not applicable.

mean, variance and coefficient of variation of kriged $\log_{10} \phi$ estimates obtained from various cross-hole test interpretations at 53,176 nodes of our auxiliary grid. The table also lists similar statistics for corresponding estimates at pilot points. As one expects, the sample mean values of kriged and pilot point estimates are very similar. The sample variances of pilot point estimates from cross-hole tests exceed those of the corresponding kriged estimates for reasons that have been explained earlier.

As noted earlier, our kriged estimates of $\log_{10} \phi$ from cross-hole tests show spatial patterns that are reminiscent of those exhibited be corresponding estimates of $\log_{10} k$. This is seen upon comparing Plate 2 with Plate 3 on the basis of the interpretation of cross-hole test PP4 and Plate 4 with Plate 5 on the basis of the joint interpretation of cross-hole tests PP4, PP5, and PP6. In both cases, regions of high and low permeability correspond quite closely to similar regions of porosity.

Kriged (dots) and pilot point (open circles) estimates of $\log_{10} \phi$ obtained from test PP4 are plotted versus corresponding estimates of $\log_{10} k$ in Figure 15. Fitting a straight line to these data by regression of $\log_{10} \phi$ on $\log_{10} k$ yields low correlation coefficients $r^2$ equal to 0.428 and 0.463 for kriged and pilot point estimates, respectively. Our hypothesis that the observed scatter can be explained by a linear trend was rejected by a standard Fisher test. The weak linear correlation may be due in part to the effect of correlated estimation errors on the scatter.

The slope of the regression line is 0.522 ± 0.004 for kriged estimates and 0.247 ± 0.174 for pilot point estimates. This is equivalent to a 1:2 linear relationship between $\log_{10} \phi$ and $\log_{10} k$ based on kriged estimates and a 1:4 linear relationship based on pilot point estimates. It is of interest to note that upon reinterpreting data from field conservative tracer tests conducted in saturated fractured rocks on various sites worldwide by different research groups, *Guimera and Carrera* [2000] obtained slopes equal to 0.28 and 0.35 for two regression lines of log effective porosity versus log permeability. These are roughly equivalent to a 1:3 linear relationship between the two parameters.

## 7. Scaling of Air Permeability and Air-Filled Porosity

Tables 6 and 7 list statistics of air permeability and air-filled porosity estimates obtained by various methods. Tables 6 and 7 reveal a steep increase in the mean values of $\log_{10} k$ and $\log_{10} \phi$ estimates with the scale of estimation. The smallest scale, of the order of 1 m, corresponds to single-hole test results and to kriged inverse estimates based on the cross-hole tests. The largest scale, of the order of a few tens of meters, corresponds to type curve results and to inverse estimates obtained from cross-hole tests upon treating the rock as being uniform. Mean values obtained by different methods at any given scale of estimation are generally comparable. Yet mean values ob-

**Table 7.** Summary Statistics of $\log_{10} \phi$ Estimated in Various Ways[a]

| Source | Data or Pilot Point Estimates | | | | Kriging Estimates | | | |
|---|---|---|---|---|---|---|---|---|
| | Sample Size | Mean | Variance | CV | Sample Size | Mean | Variance | CV |
| Cross-hole tests (type curve analyses) | | | | | | | | |
| PP4 [after *Illman and Neuman*, 2001] | 30 | −2.11 | 0.65 | −0.038 | | NA | | |
| Cross-hole tests (inverse analyses) | | | | | | | | |
| Uniform rock | | | | | | | | |
| PP4–PP8 | 128 | −1.63 | 0.24 | −0.301 | | NA | | |
| Nonuniform rock | | | | | | | | |
| PP4 | 32 | −2.92 | 2.90 | −0.584 | 53176 | −3.00 | 1.45 | −0.402 |
| PP4 | 64 | −2.79 | 1.85 | −0.487 | 53176 | −2.47 | 1.24 | −0.450 |
| PP5 | 32 | −2.84 | 2.64 | −0.573 | 53176 | −3.07 | 2.12 | −0.475 |
| PP6 | 32 | −2.11 | 0.65 | −0.382 | 53176 | −1.82 | 0.66 | −0.469 |
| PP4–PP6 | 72 | −2.74 | 1.94 | −0.510 | 53176 | −2.98 | 1.00 | −0.337 |

[a]Units of $m^3/m^3$. NA, not applicable.

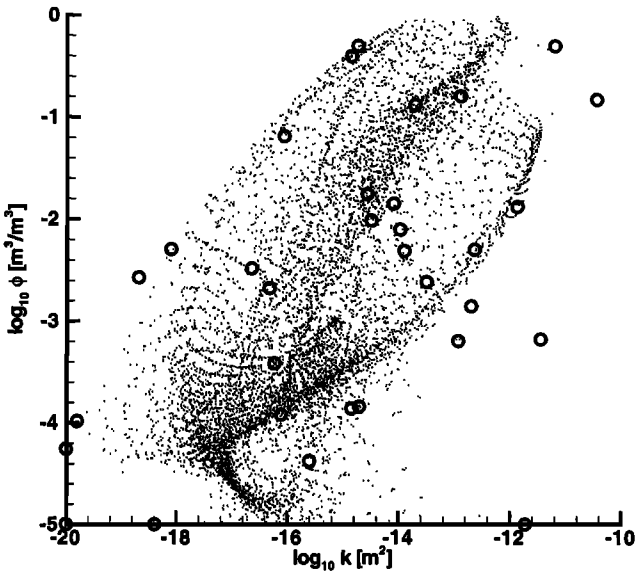


**Figure 15.** Kriged (dots) and pilot point (open circles) estimates of log air-filled porosity versus those of log air permeability based on stochastic inversion of PP4 test data.

tained under the uniform rock assumption are consistently larger than those obtained by allowing pneumatic properties to vary spatially over distances of 1 m. For example, the mean of $\log_{10} k$ estimates obtained by our inverse model from cross-hole tests PP4–PP8 when treating the rock as being uniform exceeds the mean of kriged inverse estimates from three of these tests (PP4, PP5, and PP6) exactly by 2.00. This represents a hundredfold increase in the associated values of $k$. An examination of corresponding mean $\log_{10} \phi$ estimates shows an increase in associated $\phi$ values by a factor of ~22. The mean of $\log_{10} k$ estimates obtained by our inverse model from cross-hole tests PP4–PP8 when treating the rock as being uniform exceeds the average (−15.07) of mean kriged inverse estimates from three of these tests (PP4, PP5, and PP6) by 1.38. This represents an increase in the associated values of $k$ by a factor of ~24. A similar calculation leads to an increase in $\phi$ values by a factor of ~11. It is thus clear that the estimates of both air permeability and air-filled porosity increase markedly with scale at ALRS. On the other hand, the variances of inverse $\log_{10} k$ and $\log_{10} \phi$ estimates are seen in Tables 6 and 7 to decrease with scale, as one should anticipate.

This scale effect is reflected in histograms of estimates presented in Figures 16 and 17. Histograms that represent estimates of equivalent uniform pneumatic properties over relatively large rock volumes are offset to the right relative to those that represent estimates of spatially varying properties over smaller rock volumes. Histograms belonging to the first set are narrower than those belonging to the second set. The scale effect that we observe at ALRS is clearly unrelated to the method of testing: our analysis shows consistency between single-hole and cross-hole test results. The scale effect is likewise unrelated to the method of test interpretation: our analysis shows consistency between results obtained by means of steady state analytical formulae from single-hole test data (which, in turn, compare well with transient type curve [*Illman*, 1999; *Illman and Neuman*, 2000] and inverse [*Vesselinov and Neuman*, 2001, 2000] analyses of the same data) and type curve as well as numerical inverse interpretations of cross-hole test data. Contrary to a recent suggestion by *Butler and Healey* [1998], the observed scaling behavior is not related to a skin effect: neither the single-hole [*Guzman et al.*, 1994, 1996; *Guzman and Neuman*, 1996] nor the cross-hole [*Illman et al.*, 1998; *Illman*, 1999] test results have been affected by any skin effect of consequence. Everything indicates that the pronounced permeability and porosity scale effect we observe at ALRS is real.

## 8. Conclusions

The following major conclusions can be drawn from our study:

1. It is possible to interpret cross-hole pneumatic tests in unsaturated fractured tuffs at ALRS by means of a three-dimensional numerical inverse model, which treats the rock as a locally isotropic, uniform or nonuniform continuum and air as a single mobile fluid phase. *Chen et al.* [2000] showed that the rock continuum represents primarily interconnected air-filled fractures.

2. Inverse analysis of pressure records from individual borehole monitoring intervals one at a time, while treating the

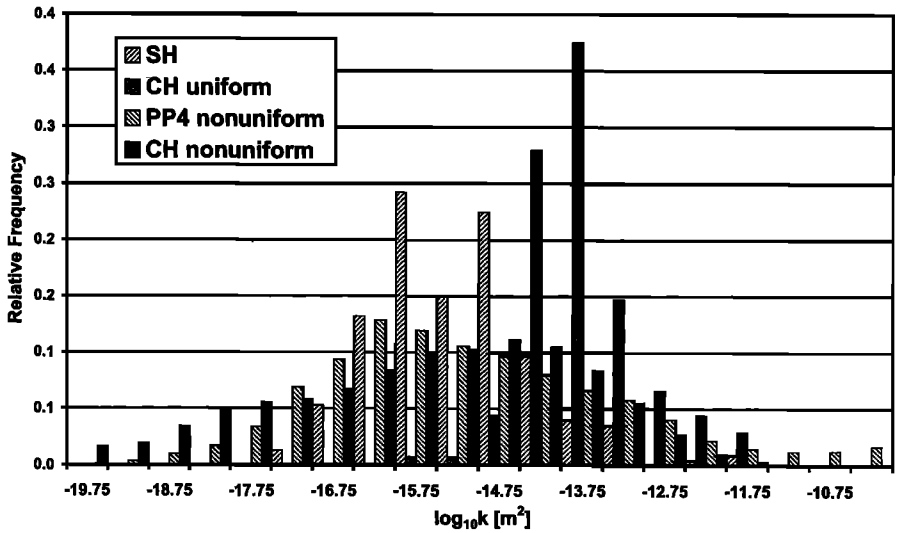


**Figure 16.** Histograms of log air permeabilities estimated by various approaches.

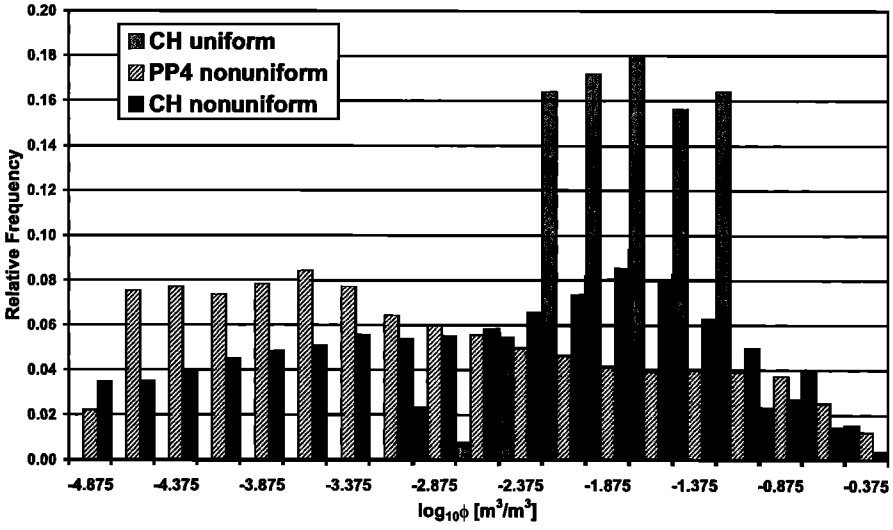


**Figure 17.** Histograms of log air-filled porosities estimated by various approaches.

rock as being spatially uniform, yields a series of equivalent air permeabilities and air-filled porosities for rock volumes having length scales ranging from meters to tens of meters, represented nominally by radius vectors extending from injection to monitoring intervals. Our equivalent parameter estimates are consistent with those obtained by *Illman and Neuman* [2001] using analytically derived type curves.

3. Inverse analysis of pressure records from multiple tests and borehole monitoring intervals simultaneously, while treating the rock as being randomly heterogeneous, yields a high-resolution kriged estimate of how air permeability and air-filled porosity, defined on grid blocks having a length scale of 1 m, vary spatially throughout the tested rock volume. It amounts to three-dimensional pneumatic "tomography" or stochastic imaging of the rock. Our kriged estimates are consistent with those obtained by *Chen et al.* [2000] from single-hole pneumatic injection tests at the site.

4. There is a weak 1:4 to 1:2 linear relationship between our estimates of log porosity and log permeability at ALRS. This is consistent with the 1:3 linear relationship found by *Guimera and Carrera* [2000] between log effective porosity and log permeability on the basis of tracer tests in saturated fractured rocks worldwide. *Chen et al.* [2000] have shown that log air permeability correlates poorly with fracture densities at ALRS.

5. There is a very pronounced scale effect in air permeability and air-filled porosity at ALRS. Because there is consistency between single-hole and cross-hole test results, the scale effect is unrelated to the method of testing. Because there is consistency between results obtained by means of diverse steady state and transient, analytical, and numerical methods of test interpretation, the scale effect is unrelated to the method of interpretation. As neither the single-hole nor the cross-hole test results have been affected by any skin effect of consequence, the scale effect is unrelated to skin phenomena. The observed scale effect at ALRS appears to be real.

6. The novelty of our inverse methodology was highlighted in paper 1. Whereas the principle behind this methodology is general, its application to cross-hole tests in unsaturated fractured tuffs at the Apache Leap Research Site yields parameter estimates that are unique to this site. They should, however, be indicative of pneumatic properties of partially welded unsaturated fractured tuffs at other locations, such as Yucca Mountain in Nevada.

## Appendix A: Effect of Linearization on Evaluation of $k$ and $\phi$

Consider (1)–(4) in paper 1. Disregarding gravity and sources leads to the following equations for isothermal airflow in a uniform medium:

$$\nabla \cdot (p\nabla p) = \frac{\mu\phi}{k}\frac{\partial p}{\partial t} \tag{A1}$$

subject to initial conditions given by (2) in paper 1 and the generalized boundary condition

$$\frac{kM_m}{\mu ZRT}[(p\nabla p)\cdot \mathbf{n}] = v(p_f - p) + q_b \tag{A2}$$

along $\Gamma$. All terms in these equations are defined in paper 1. Note that (A2) is in terms of mass flux.

*Illman and Neuman* [2001] use a so-called $p$-based linearization

$$\nabla^2 p = c_1\frac{\partial p}{\partial t} \tag{A3}$$

$$c_2[\nabla p \cdot \mathbf{n}] = v(p_f - p) + q_b \tag{A4}$$

along $\Gamma$, where

$$c_1 = \mu\phi\bar{C}_a/k \tag{A5}$$

$$c_2 = kM_m/\mu ZRT\bar{C}_a, \tag{A6}$$

with $\bar{C}_a$ being a constant air compressibility given by $1/\bar{p}$, $\bar{p}$ is a reference air pressure that *Illman and Neuman* [2001] set equal to the initial (barometric) pressure $p_0$.

*Illman and Neuman* [2001] determine $c_1$ and $c_2$ by matching recorded pressures to type curves based on these linearized equations, then evaluate $k$ from (A6) and $\phi$ from (A5). Thus, for given $c_1$ and $c_2$, $k$ depends on $\bar{C}_a$ but $\phi$ does not. As $\bar{C}_a$ depends on barometric pressure, so does $k$.

A question arises as to how would a constant increment $p_s$ in

the barometric pressure $p_0$, given $c_1$ and $c_2$, affect the evaluation of $k$. Clearly, if $\hat{p}$ is the solution of the linearized equations for $p_0$, the solution for a barometric pressure $p_0 + p_s$ is $\hat{p} + p_s$. It follows that $\bar{C}_a$ is given by $1/(p_{\text{atm}} + p_s)$ and so (A6) would yield a value of $k$ that differs by factor of $p_0/(p_0 + p_s)$ from that one would have calculated with $p_0$. On the other hand, $\phi$ would remain unaffected by $p_s$. It follows that $k$ estimates corresponding to a barometric pressure of 100 kPa could be made to correspond to a barometric pressure of 87 kPa ($p_0 = 100, p_s = -13$) simply upon multiplying them by 100/87 = 1.16.

*Illman and Neuman* [2001] looked at the same question by considering a volume-based version of (A2), obtained upon dividing (A2) by the air density $\rho$,

$$\frac{k}{\mu}(\nabla p)\cdot\mathbf{n} = v'(p_f - p) + q_b' \qquad \text{(A7)}$$

along $\Gamma$, where $v'$ $[L^2T/M]$ and $q_b'$ $[L/T]$ are volume-based versions of $v$ and $q_b$ ($v' = v/\rho$, $q_b' = q_b/\rho$). This is valid provided the density and pressure of injected air are the same as in the rock near the injection interval. In examining the question how does a constant increment $p_s$ in the barometric pressure $p_0$, given $c_1$ and $c_2$, affect the evaluation of pneumatic parameters, *Illman and Neuman* [2001] neglected to consider the effect of $p_s$ on $q_b'$. They therefore concluded that $p_s$ would affect the evaluation of $\phi$ (by a factor of $(p_{\text{atm}} + p_s)/p_{\text{atm}}$) but not of $k$.

Values of $k$ calculated by *Illman and Neuman* [2001] for cross-hole test PP4 exceed those we estimate numerically for a uniform medium by a factor of ~1.7. Yet late time pressure in the injection interval during this test exceeded their reference pressure $p_0$ by a factor of ~2. Had Illman and Neuman reduced $q_b'$ in (A7) by a similar factor, they would have obtained $k$ estimates that are about half as large as those they list and so closer to our estimates.

To address the question how a constant change $p_s$ in reference pressure would affect the values of pneumatic parameters that we compute by means of the original nonlinear airflow equations, consider replacing $p$ in (A1) and (A2) by $p + p_s$. Whereas $\nabla p$ and $\partial p/\partial t$ would remain unaffected, $p\nabla p \neq (p + p_s)\nabla(p + p_s)$ in (A1) and (A2) would change. This would impact the evaluation of $k$ but not $\phi$, just as in the linearized case.

**Acknowledgments.** This research was supported in part by the U.S. Nuclear Regulatory Commission under contracts NRC-04-95-038 and NRC-04-97-056. Some of the simulations and inverse modeling were conducted during a summer internship of the senior author with the Geoanalysis Group at Los Alamos National Laboratory. We are grateful to George A. Zyvoloski for his help in the implementation of FEHM and to Carl W. Gable for his assistance in the use of X3D. We thank the anonymous reviewers for their constructive and insightful comments.

## References

Butler, J. J., and J. Healey, Relationship between pumping-test and slug-test parameters: Scale effect or artifact?, *Ground Water*, *36*, 305–313, 1998.

Chen, G., W. A. Illman, D. L. Thompson, V. V. Vesselinov, and S. P. Neuman, Geostatistical, type-curve and inverse analyses of pneumatic injection tests in unsaturated fractured tuffs at the Apache Leap Research Site near Superior, Arizona, in *Dynamics of Fluids in Fractured Rocks, Geophys. Monogr. Ser.*, vol. 122, edited by B. Faybishenko, P. A. Witherspoon, and S. M. Benson, pp. 73–98, AGU, Washington, D. C., 2000.

Doherty, J., L. Brebber, and P. Whyte, *PEST: Model Independent Parameter Estimation*, Watermark Comput., Brisbane, Queensland, Australia, 1994.

Guimera, J., and J. Carrera, A comparison of hydraulic and transport parameters measured in low permeability fractured media, *J. Contam. Hydrol.*, *41*, 261–281, 2000.

Guzman, A. G., and S. P. Neuman, Field air injection experiments, in *Apache Leap Tuff INTERVAL Experiments*, edited by T. C. Rasmussen et al., *Rep. NUREG/CR-6096*, pp. 52–94, U.S. Nucl. Regul. Comm., Washington, D. C., 1996.

Guzman, A. G., S. P. Neuman, C. F. Lohrstorfer, and R. L. Bassett, Field hydraulic, pneumatic and tracer tests: Phase I, in *Validation Studies for Assessing Unsaturated Flow and Transport Through Fractured Rock*, edited by R. L. Bassett et al., *Rep. NUREG/CR-6203*, U.S. Nucl. Regul. Comm., Washington, D. C., 1994.

Guzman, A. G., A. M. Geddis, M. J. Henrich, C. F. Lohrstorfer, and S. P. Neuman, Summary of air permeability data from single-hole injection test in unsaturated fractured tuff at the Apache Leap Research Site: Results of steady-state test interpretation, *Rep. NUREG/CR-6360*, U.S. Nucl. Regul. Comm., Washington, D. C., 1996.

Illman, W. A., Single- and cross-hole pneumatic injection tests in unsaturated fractured tuffs at the Apache Leap Research Site near Superior, Arizona, Ph.D. thesis, Univ. of Ariz., Tucson, 1999.

Illman, W. A., and S. P. Neuman, Type-curve interpretation of multirate single-hole pneumatic injection tests in unsaturated fractured rock, *Ground Water*, *38*, 899–911, 2000.

Illman, W. A., and S. P. Neuman, Type curve interpretation of a cross-hole pneumatic injection test in unsaturated fractured tuff, *Water Resour. Res.*, *37*, 583–604, 2001.

Illman, W. A., D. L. Thompson, V. V. Vesselinov, G. Chen, and S. P. Neuman, Single- and cross-hole pneumatic tests in unsaturated fractured tuffs at the Apache Leap Research Site: Phenomenology, spatial variability, connectivity and scale, *Rep. NUREG/CR-5559*, U.S. Nucl. Regul. Comm., Washington, D. C., 1998.

Neuman, S. P., Stochastic continuum representation of fractured rock permeability as an alternative to the REV and fracture network concepts, *Proc. U.S. Symp. Rock Mech.*, *28*, 533–561, 1987.

Pebesma, E., and C. Wesseling, GSTAT, a program for geostatistical modelling, prediction and simulation, *Comput. Geosci.*, *24*, 17–31, 1998.

Trease, H. E., D. George, C. W. Gable, J. Fowler, A. Kuprat, and A. Khamyaseh, The X3D grid generation system, in *5th International Conference on Numerical Grid Generation in Computational Fluid Dynamics and Related Fields*, edited by B. K. Soni et al., pp. 239–244, Eng. Res. Cent., Miss. State Univ. Press, Mississippi State, 1996.

Vesselinov, V. V., Numerical inverse interpretation of pneumatic tests in unsaturated fractured tuffs at the Apache Leap Research Site, Ph.D. dissertation, Univ. of Ariz., Tucson, 2000.

Vesselinov, V. V., and S. P. Neuman, Numerical inverse interpretation of multistep transient single-hole pneumatic tests in unsaturated fractured tuffs at the Apache Leap Research Site, *348*, Geol. Soc. Am. *Spec. Pap.*, 175–190, 2000.

Vesselinov, V. V., and S. P. Neuman, Numerical inverse interpretation of single-hole pneumatic tests in unsaturated fractured tuff at the Apache Leap Research Site, *Ground Water*, *36*, 685–695, 2001.

Vesselinov, V. V., S. P. Neuman, and W. A. Illman, Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff, 1, Methodology and borehole effects, *Water Resour. Res.*, this issue.

Zyvoloski, G. A., B. A. Robinson, Z. V. Dash, and L. L. Trease, Summary of the models and methods for the FEHM application: A finite-element heat- and mass-transfer code, *Rep. LA-13307-MS*, Los Alamos Natl. Lab., Los Alamos, N. M., 1997.

W. A. Illman, Department of Geosciences, University of Iowa, 121 Trowbridge Hall, Iowa City, IA 52242, USA. (walter-illman@uiwa.edu)

S. P. Neuman, Department of Hydrology and Water Resources, University of Arizona, Bldg. 11, P.O. Box 210011, Tucson, AZ 85721-0011, USA. (neuman@hwr.arizona.edu)

V. V. Vesselinov, Hydrology, Geochemistry, and Geology Group, Los Alamos National Laboratory, EES-6, MS T003, Los Alamos, NM 87545, USA. (vvv@lanl.gov)

(Received December 6, 2000; revised May 15, 2001; accepted June 1, 2001.)