LBNL-39620

Design and Analysis of a Well Test for Determining Two-Phase Hydraulic Properties

Stefan Finsterle and Karsten Pruess

Lawrence Berkeley National Laboratory
Earth Sciences Division
University of California
Berkeley, CA 94720

December 1996

This work was supported by a grant from the Swiss National Cooperative for the Disposal of Radioactive Waste (Nagra), Wettingen, Switzerland, through the U.S. Department of Energy Contract No. DE-AC03-76SF00098.

Design and analysis of a well test for determining two-phase hydraulic properties

Stefan Finsterle and Karsten Pruess

Earth Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley

Abstract. *This paper describes the design and analysis of a well test for determining hydraulic properties of a low permeability, low porosity formation that potentially contains a small amount of free methane. Estimation of gas-related parameters in such formations is difficult using standard pumping tests because (i) pressures and flow rates may fluctuate as a consequence of gas exsolution during the test, (ii) data may not allow discriminating among alternative conceptual models, making the parameter estimates ambiguous, and (iii) the key parameters of interest are highly correlated. In this study we adopt an inverse modeling perspective to examine a test sequence that can be appended to a standard hydraulic testing program. The design calculations show that a series of water and gas injections can significantly reduce parameter correlations, thus decreasing estimation uncertainty. Moreover, the extended test sequence enhances the possibility of identifying the model describing relative permeabilities and capillary pressures. A prerequisite for a successful inversion is that data of high accuracy are collected under controlled test conditions. We discuss the assumptions and limitations of the procedure, and propose some recommendations for testing. Finally, we analyze pressure and gas flow data from a well test performed by the Swiss National Cooperative for the Disposal of Radioactive Waste (Nagra) that partially followed the proposed test sequence.*

1. Introduction

One of the main purposes of flow and transport modeling is to predict the future behavior of a subsurface flow system. The reliability of such model predictions, however, has to be questioned because of the uncertainties in the conceptual model and the input parameters. In most cases a flaw in the model structure leads to a considerable prediction error [*Morgan and Henrion*, 1990; *Usunoff et al.*, 1992; *James and Oldenburg*, 1996], making the choice of the conceptual model the most important step in model development. The second largest source of uncertainty stems from insufficient knowledge about the model parameters and their variability [*Dettinger and Wilson*, 1981; *Wagner and Gorelick*, 1987; *Woldt et al.*, 1992]. One way to obtain parameter estimates and their uncertainties is to calibrate the model against observations [*Yeh*, 1986]. Since parameters estimated using inverse modeling can be considered optimal for the given model, an unacceptably large fitting residual may indicate the existence of a model structure error [*Carrera and Neuman*, 1986]. Unfortunately, a good match to the data does not prove that the conceptual model is correct [*Finsterle and Persoff*, 1996]. It could be that the data are not sensitive with respect to certain aspects of the model. If these aspects become relevant during the subsequent model application, the predictions may be erroneous despite successful history matching and small parameter uncertainties. This problem is reflected in the statement that all parameter values estimated by data inversion are strictly model-related [*Carrera and Neuman*, 1986]. This fact can be an advantage when the key aspects of a model are not changed between history matching and prediction runs. In these cases, inverse modeling provides effective parameters which can be considered optimal for the given application. However, caution has to be exercised whenever parameters are transferred to a modified or new model. It should be noted that this difficulty is intrinsic to all estimation techniques, including direct measurement of parameters.

The objective of this paper is to demonstrate how inverse modeling techniques can be used to optimize the design of an experiment. We will deduce criteria for choosing a superior test design from a set of competing alternatives. The framework for such an optimization is given by the overall purpose of the experiment, which determines the degree of acceptable parameter uncertainty. We also address the issue of model identification as the primary source of systematic parameter and prediction errors.

We first discuss the general idea of performing inversions as a central part of design calculations. We then introduce the criteria and performance measures used to examine competing test designs. The simulation model is briefly described, after which we apply the approach to a well test designed for the detection of small gas contents in a tight formation.

2. Inverse Modeling and Test Design

Finding an optimal test design is an iterative process which involves technical and non-technical issues and objectives. Weighting diverse goals is usually based on management decisions which are difficult to incorporate into a mathematical optimization model. Nevertheless, there are certain aspects of test design that can be subjected to rigorous analysis. Reviews of experimental design procedures are given by *Steinberg and Hunter* [1984] and *Sun and Yeh* [1990]. *Knopman and Voss* [1988, 1989] and *Knopman et al.* [1991] have described a sampling design methodology for solute transport problems based on three objectives: model discrimination, parameter estimation, and cost minimization.

Their objective function for model discrimination considers prediction divergence as a desirable attribute of a good design. Model discrimination was also addressed by *Carrera and Neuman* [1986] and *Usunoff et al*. [1992], relying on statistical criteria such as the estimated error variance or the Kashyap criterion [*Kashyap*, 1982]. The parameter estimation objective is usually based on some measure of estimation uncertainty, an approximation of which can be calculated from the sensitivity matrix. The latter is also needed for the solution of the nonlinear regression problem, and contains useful information about the value of individual data points. In this paper we take a similar approach, focusing on the underlying multiphase flow problem and the discussion of a few performance measures.

We presume that the objective of an experiment is to determine certain parameter values which characterize the hydrogeologic properties (e.g. permeability) or the in-situ conditions (e.g., gas content) of the formation. Note that the set of parameters to be estimated as well as the required accuracy are deduced from the overall objective of a project. It is the acceptable level of prediction uncertainty, and the sensitivity of these predictions to various model parameters, that tells us which parameters are relevant and what upper limit of the estimation error shall not be exceeded. For example, the gas entry pressure of a saturated formation is considered to be an important parameter for the prediction of pressure buildup in a repository for gas-generating radioactive wastes located in the saturated zone. Therefore, the entry pressure as a parameter of the capillary pressure function has to be determined with a relatively high degree of accuracy in order to limit prediction uncertainty. On the other hand, the entry pressure may not significantly affect the system behavior of a repository that is situated in the unsaturated zone, where other parameters may be of greater concern. Differences in overall objectives may greatly influence the design of an experiment. In this paper we assume that the test objectives are given, along with a list of parameters of interest and their acceptable estimation uncertainty.

The data collected during an experiment measure the response of the system to the applied perturbation. If the system response is sensitive to the parameters of interest, inverse modeling provides a means to estimate the parameter values that are most likely to have produced the observed system response. Recall that those parameters refer to the conceptual model that was set up to mimic the experiment. In case of a mismatch between the observed data and the corresponding calculated system response, a detailed residual analysis may reveal some aspects of the model that need to be modified.

There are two elements of experimental design which can be optimized by inverse-modeling design calculations. First, we can study different perturbations to determine which kind of experiment would be most suitable to identify the parameters of interest. We will refer to different experimental layouts as alternative test designs. Secondly, the type, amount, location, and accuracy of the observations can be varied to improve parameter identifiability. The information content of an individual data point can be evaluated by a standard sensitivity analysis. However, the issue of parameter correlation can only be addressed by taking an inverse perspective. Next we need criteria to judge the performance of competing test designs. The main criterion will be the uncertainty of the estimated parameters, but we can also look at the ability of an experiment to discriminate between alternative conceptual models. Both criteria require performing inverse modeling runs rather than doing a standard sensitivity analysis. Our approach to design calculations mirrors the procedure that will be applied for the subsequent data analysis and involves the following steps (see Figure 1):

1. Define a conceptual model that most likely represents the system under study. Conceive a test design, i.e., choose the sequence of test events, the type of data to be collected, the location and accuracy of sensors, etc.

2. Generate synthetic data for all potential observation points through forward modeling of the test sequence; random noise may be added to the synthetic data.

3. Solve the inverse problem for all unknown or uncertain parameters.

4. Analyze the Jacobian matrix which provides information regarding the sensitivity of each observation with respect to each parameter. Revise the test design to increase sensitivity.

5. Analyze the covariance matrix of the estimated parameters. Optimize the test design to reduce estimation uncertainty and parameter correlation.

6. Change the model structure, and again try to fit the synthetic data. If they can be matched equally well regardless of the model being used, then the test design does not produce selective data, i.e., an erroneous conceptual model is not rejected, and the resulting parameter estimates will be biased. Revise the test design to produce selective data.

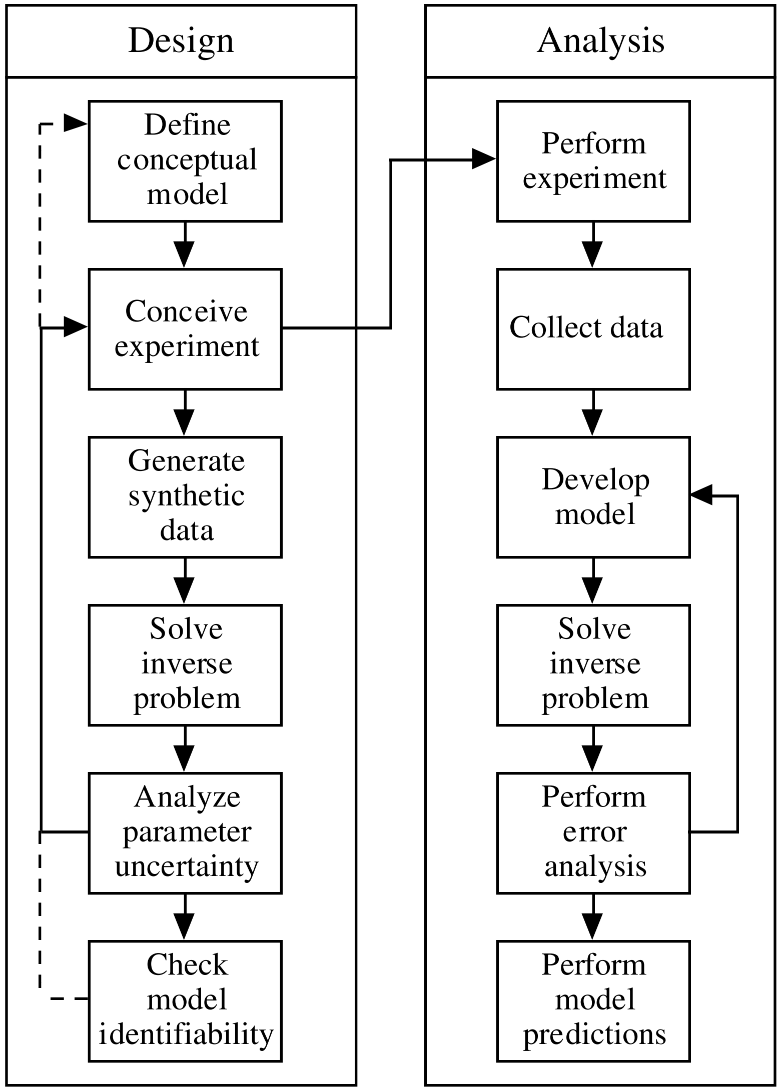

Figure 1. Procedure of test design and data analysis.

In the remainder of this section we elaborate on steps 4 through 6, above. The calculated system response (e.g., pressure at a certain point in space and time) will be referred to as z_i; parameters are designated with p_k. First of all, estimates of measurement errors have to be specified. Potential measurement errors, assumed to be normally distributed about a mean, can be described by a covariance matrix $\mathbf{C}_{zz}$. Note that only the relative magnitude of the elements of $\mathbf{C}_{zz}$ will influence the values of the estimated parameters. We therefore introduce a dimensionless factor σ_0^2 which is termed the prior error variance, and a positive definite matrix $\mathbf{V}_{zz}$, where $\mathbf{V}_{zz}^{-1}$ will be used as a weighting matrix:

$$\mathbf{C}_{zz} = \sigma_0^2 \mathbf{V}_{zz} \tag{1}$$

While σ_0^2 can assume any positive number, it is convenient to set $\sigma_0^2 = 1$, i.e., the weighting matrix is the inverse of the covariance matrix. If actual measurements were available, the estimated error variance s_0^2 after matching the data would be given by

$$s_0^2 = \frac{\mathbf{r}^{\mathrm{T}} \mathbf{V}_{zz}^{-1} \mathbf{r}}{m - n} \tag{2}$$

where $\mathbf{r}$ is the vector of residuals, comprised of the differences between observed and calculated system response, m is the number of observations, and n is the number of parameters. If a perfect fit is obtained using a large number of calibration points, the standard deviation of the residuals, s_0^2, approaches the measurement error. Since s_0^2 is a random variable, it can be statistically tested against the prior error variance σ_0^2. If the ratio s_0^2 / σ_0^2 is significantly larger than 1 based on a Fisher model test [Cooley et al., 1986], this indicates either that the model is inappropriate to reproduce the data (wrong conceptual model), or that the magnitude of the measurement errors reflected by the matrix $\mathbf{C}_{zz}$ was underestimated. Note that when design calculations are performed, no data are available yet and thus no residuals can be calculated. However, expectations regarding the residuals can be expressed through matrix $\mathbf{C}_{zz}$, allowing σ_0^2 to replace s_0^2 in the uncertainty analysis described below.

The covariance matrix of the estimated parameters $\mathbf{C}_{pp}$ can be approximated by

$$\mathbf{C}_{pp} = s_0^2 \left(\mathbf{J}^{\mathrm{T}} \mathbf{V}_{zz}^{-1} \mathbf{J} \right)^{-1} \tag{3}$$

where $\mathbf{J}$ is the Jacobian matrix of dimension $m \times n$ with elements $J_{ik} = \partial z_i / \partial p_k$. Note that (3) is a linear approximation of the actual confidence region. Methods to obtain improved estimates of $\mathbf{C}_{pp}$ for highly nonlinear models are described by *Carrera* [1984], *Hamilton and Watts* [1985], and *Finsterle and Pruess* [1995]. If actual data were analyzed, the Jacobian would be evaluated at the optimum parameter set; for design calculations, p_k has to be replaced by the expected parameter value, i.e. the value which is believed to best represent the actual, albeit unknown system. It is obvious that due to the nonlinearities in multiphase flow modeling, the results of the design calculations are local in the sense that they are valid only for the prior estimate of the unknown parameters, i.e., they may change considerably if the actual field conditions are different from the expected ones. *Casman et al*. [1988] address this issue by evaluating an efficiency function which measures the ability of an experimental design to reduce the variance of the parameter estimates.

In all design evaluations discussed in the literature, the interpretation of the covariance matrix $\mathbf{C}_{\mathrm{pp}}$ provides the key criteria based on which the experimental design can be improved. First we note that $\mathbf{C}_{\mathrm{pp}}$ is directly proportional to the overall goodness-of-fit expressed by s_0^2 or - in the case of design calculations - the expectation thereof. The latter can easily be modified by changing σ_0^2. The diagonal elements of $\mathbf{C}_{\mathrm{pp}}$ contain the variances $\sigma_{p_k}^2$ of the estimated parameters p_k. The test design should be optimized primarily with respect to this measure, i.e. the test sequence which yields smaller $\sigma_{p_k}^2$ performs better than its competing alternative, providing for more accurate estimates. There are three scalar measures of $\mathbf{C}_{\mathrm{pp}}$ one might use as design criteria. Taking the determinant of $\mathbf{C}_{\mathrm{pp}}$ yields the so-called *D*-optimality objective for design evaluation. The local *D*-optimality solution minimizes the area of the joint confidence region around the parameter estimates. Alternatively, *A*-optimality consists of minimizing the trace of $\mathbf{C}_{\mathrm{pp}}$, and *E*-optimality seeks minimization of the maximum eigenvalue of $\mathbf{C}_{\mathrm{pp}}$ [*Steinberg and Hunter*, 1984]. If vector $\mathbf{p}$ contains parameters of different types and orders of magnitude, matrix $\mathbf{C}_{\mathrm{pp}}$ should be appropriately scaled before evaluating the optimality criteria.

Next we briefly discuss the impact of correlations on the estimation error. Correlations among parameters can be described as the combined impact of parameter changes on the system behavior. For example, if two parameters are negatively correlated, a similar system response is obtained by concurrently increasing one and decreasing the other parameter. Even though certain pairs of parameters may exhibit preferential correlation structures, correlations are not invariable for a given parameter combination. They obviously depend on the data available, and also on the number of simultaneously estimated parameters, since indirect correlations may overwhelm the direct correlations [*Finsterle and Pruess*, 1995]. If correlations exist, the uncertainty of one parameter does affect the uncertainty of the other parameter. The diagonal elements of matrix $\mathbf{C}_{\mathrm{pp}}$, which are the variances from the joint probability density function, account for this fact. They have to be distinguished from the marginal standard deviation $\sigma_{p_k}{}^*$ which measures the uncertainty of a parameter assuming that all the other parameters are either exactly known or uncorrelated. It is obvious that the standard deviation from the marginal distribution is always smaller than the one from the joint probability density function. The ratio

$$\chi_k = \frac{\sigma_{p_k}{}^*}{\sigma_{p_k}} \tag{4}$$

can therefore be regarded as a measure of overall parameter correlation, i.e., of how independently parameter k can be estimated. Small values of χ_k usually indicate that the uncertainty σ_{p_k} of a parameter can be further reduced by lowering its correlation to other parameters. Taking into account the correlations between the parameters is one of the key advantages of the proposed inversion procedure over conventional sensitivity analysis. If parameter correlations are ignored during the design stage of an experiment, the problem of parameter ambiguity is not properly addressed, and the expected accuracy of the estimates is likely to be overestimated.

So far we have discussed the criteria for judging the performance of a test design. No guidelines have been given regarding the optimization process itself, i.e. how the test design could be improved to actually reduce σ_{p_k}. No general recommendation can be given. However, the Jacobian matrix **J** can be used to identify location, time segment, and data type which most likely carry information about the model parameters. The Jacobian contains the sensitivity of each data point with respect to each parameter. We propose to scale the coefficients of the Jacobian by the expected variation of the parameter and by the inverse of the standard deviation of the observations:

$$\psi_{ik} = \frac{\partial z_i}{\partial p_k} \cdot \frac{\sigma_{p_k}}{\sigma_{z_i}} \tag{5}$$

With this definition, the contribution of each potential data point to the solution of the inverse problem at hand can be evaluated by calculating an integral measure ζ_i ,

$$\zeta_i = \sum_{k=1}^{n} \left|\psi_{ik}\right| \tag{6}$$

For example, simulated data points with very low ζ_i-values can be discarded without loss of information, i.e. the corresponding sensor should be moved in space, activated during a different time segment, or made more accurate until a higher ζ_i-value is realized. This will improve the overall sensitivity and thus reduce the estimation error. Similarly, additional data should be taken in regions with high ζ_i-values. However, if sampling points are set too close to each other in space and time, only a minor improvement will be achieved due to high correlations. This will be reflected in a small reduction of the determinant of $\mathbf{C}_{pp}$ despite a large increase of the sum of all ζ_i-values.

One might also compare the overall parameter sensitivity ϖ_k:

$$\varpi_k = \sum_{i=1}^{m} |\psi_{ik}| \tag{7}$$

to identify the most sensitive parameters as well as those for which no sensitive data are available. Adding new data may help improve the total sensitivity of a parameter which eventually makes possible the estimation of its value with an acceptable degree of uncertainty. The measures ζ_i and ϖ_k are relative quantities, i.e. they scale with our assumption regarding the accuracy of the sensors, σ_{z_i}, and the expected parameter variation, σ_{p_k}. However, they point towards aspects of the test configuration that can be modified to improve the overall design.

We finally address the issue of model identification. As a general rule, an experiment should be configured such that the salient features of the system and its model representation are revealed. In other words, a competing model should fail if it lacks some of the characteristics which are pertinent to the system. The relative performance of different conceptual models can be measured by one of the identification criteria discussed by *Carrera* [1984]. The nonlinearity of the multiphase flow model studied here leads to a dependence of most of the criteria discussed in the literature on the unknown parameters. Furthermore, the various simplifications made in the design stage of an experiment preclude us from using highly sophisticated model discrimination and design criteria (for a discussion see *Casman et al*. [1988] and *Sun and Yeh* [1990]). We therefore simply rely on the ratio of the estimated error variances (Eq. 2) for two models to measure the relative ability of models to match the synthetic data. A conservative test of the hypothesis that the estimated error variance of model A is significantly smaller than the one of model B is given by the critical value of the F-distribution on a certain significance level (1- α) [*Cooley et al.*, 1986]:

$$\frac{s_{0B}^2}{s_{0A}^2} > F_{(m-n)_B,(n-m)_A,1-\alpha} \tag{8}$$

In a design study, the model having generated the synthetic data is perfectly known; it should be the only one identified among the set of alternative models, provided that the data is selective with respect to changes of the model structure. If the "true" model does not realize a significantly lower s_0^2-value than the competing alternatives, the test design has to be changed to produce more selective data.

Due to incomplete parametrization of a test design, and a great deal of subjective judgment in the definition of the design criteria, we do not follow a fully formalized approach by actually minimizing a design objective. Instead, we simply evaluate and compare the performance of a limited number of designs. We propose to conduct design calculations by solving the inverse problem for synthetically generated data, and by examining the potential estimation error. The test design can be improved towards smaller standard deviations of the parameters of interest, taking advantage of the information provided by the sensitivity matrix. The ability of the design to discriminate among competing model structures can also be scrutinized. While the proposed procedure exhibits some of the shortcomings inherent in any design calculation, it has substantial advantages over a standard sensitivity analysis. Taking an inverse approach allows one to address the issues of parameter uncertainty, uniqueness, instability, and correlations among the parameters, thus reducing the risk to collect data which do not contain conclusive information regarding the parameters of interest.

3. Forward and Inverse Modeling Approach

The approach discussed in this paper makes use of a multiphase flow simulator in combination with a flexible, robust and efficient inversion technique. First, we need a computer program that allows us to simulate the proposed experiment, taking into account all the relevant processes affecting the observable system response. We use the TOUGH2 simulator [*Pruess*, 1987, 1991] to model a system with two mobile phases β (β = g: gas; β = l: liquid), and two components κ ($\kappa = CH_4$: methane; $\kappa = w$: water). The governing mass-balance equation for each component can be written in the following integral form [*Pruess and Narasimhan*, 1985]:

$$\frac{\partial}{\partial t}\int_V M^\kappa dv = \int_\Gamma \mathbf{F}^\kappa \cdot \mathbf{n} d\Gamma + \int_V q^\kappa dv \tag{9}$$

The integration is over an arbitrary subdomain V of the flow system which is bounded by the closed surface Γ with inward normal vector $\mathbf{n}$. M^κ is the mass accumulation term for component κ, $\mathbf{F}^\kappa$ is the mass flux term, and q^κ is a term representing sinks and sources. The mass accumulation term is:

$$M^\kappa = \phi \cdot \sum_{\beta=l,g} S_\beta \cdot \rho_\beta \cdot X_\beta^\kappa \tag{10}$$

where ϕ is porosity, S_β is phase saturation, ρ_β is density of phase β, and X_β^κ is the mass fraction of component κ in phase β. Thus, M^κ is the total mass of component κ present per unit volume. The mass flux terms contain a sum over the two phases,

$$\mathbf{F}^\kappa = \sum_{\beta=l,g} \mathbf{F}_\beta \cdot X_\beta^\kappa \tag{11}$$

where each phase flows in response to pressure and gravitational forces according to the multiphase extension of Darcy's law:

$$\mathbf{F}_\beta = -k \frac{k_{r\beta}}{\mu_\beta} \rho_\beta \left(\nabla p_\beta - \rho_\beta \cdot \mathbf{g} \right) \tag{12}$$

Here, k denotes absolute permeability, $k_{r\beta}$ is relative permeability of phase β, μ_β is dynamic viscosity, p_β is the pressure in phase β, and $\mathbf{g}$ is gravitational acceleration. The thermophysical properties of water are calculated from the steam table equations as given by the *International Formulation Committee* [1963]. The gas phase is treated as an ideal gas, and additivity is assumed for methane and vapor partial pressures.

Different models are used in this study to calculate relative permeabilities $k_{r\beta}$ and capillary pressure p_c as a function of liquid saturation S_l. Note that the functional form is part of the conceptual model, whereas the individual parameters entering those functions can be subjected to the estimation process. Some of the parameters used in the models have similar physical meaning, others are only related by analogy. In order to simplify the discussion, we will refer to some of them using the names given in Table 1.

The first model (LI) consists of linear functions:

$$p_c = -p_e + \left(p_{\max} - p_e \right)\left(S_{ec} - 1 \right) \tag{13}$$

$$k_{rl} = S_{ek} \tag{14}$$

$$k_{rg} = 1 - S_{ek} \tag{15}$$

where the effective saturations for the capillary pressure and relative permeabilities are given, respectively, by

$$S_{ec} = \frac{S_l - S_{lr}}{1 - S_{lr}} \qquad \left(S_{lr} < S_l \le 1\right) \tag{16}$$

$$S_{ek} = \frac{S_l - S_{lr}}{1 - S_{lr} - S_{gr}} \qquad \left(S_{lr} < S_l \le 1 - S_{gr}\right) \tag{17}$$

With $m = 1 - 1/n$, the van Genuchten model (VG) can be written as [*Luckner et al*., 1989]:

$$p_c = -\frac{1}{\alpha}\left[\left(S_{ec}\right)^{-1/m} - 1\right]^{1/n} \tag{18}$$

$$k_{rl} = S_{ek}{}^{1/2}\left[1 - \left(1 - S_{ek}{}^{1/m}\right)^m\right]^2 \tag{19}$$

$$k_{rg} = \left(1 - S_{ek}\right)^{1/3}\left[1 - S_{ek}{}^{1/m}\right]^{2m} \tag{20}$$

The Brooks-Corey model (BC) is given by [*Luckner et al*., 1989]:

$$p_c = -p_e\left(S_{ec}\right)^{-1/\lambda} \tag{21}$$

$$k_{rl} = S_{ek}{}^{\frac{2-3\lambda}{\lambda}} \tag{22}$$

$$k_{rg} = \left(1 - S_{ek}\right)^2\left(1 - S_{ek}{}^{\frac{2+\lambda}{\lambda}}\right) \tag{23}$$

The model referred to as Grant's curve is identical with the Brooks-Corey model except for its higher gas relative permeability which is given by:

$$k_{rg} = 1 - k_{rl} \tag{24}$$

Table 1. Parameters of Characteristic Curves

Symbol	Units	Generic Parameter Name
S_{lr}	-	residual liquid saturation
S_{gr}	-	residual gas saturation
λ, n	-	pore size distribution index
p_e, $1/\alpha$	Pa	gas entry pressure
$p_{\max}$	Pa	capillary pressure for $S_l \le S_{lr}$

In addition to phase interference, TOUGH2 also models interphase mass transfer assuming local chemical and thermal equilibrium.

The inverse problem of groundwater modeling is described, for example, in *Carrera and Neuman* [1986], who take the view of maximum likelihood estimation, and is comprehensively reviewed by *Yeh* [1986], *Carrera* [1988], and *McLaughlin and Townley* [1996]. Parameter estimation for solute transport and unsaturated flow models is reviewed by *Kool et al.* [1987]. We use the ITOUGH2 code [*Finsterle*, 1993; *Finsterle and Pruess*, 1995] for the solution of the inverse problem. ITOUGH2 allows estimation of any TOUGH2 input parameter including initial and boundary conditions, based on any type of sensitive observation for which a corresponding TOUGH2 output can be calculated. The inverse problem is solved by minimizing the weighted least-squares objective function

$$Z(\mathbf{p}) = \mathbf{r}^T \mathbf{V}_{zz}^{-1} \mathbf{r} \tag{25}$$

The Levenberg-Marquardt modification of the Gauss-Newton algorithm [*Levenberg*, 1944; *Marquardt*, 1963] has been found to be the most robust for our purposes. The basic idea of this method is to move in the parameter space along the steepest descent direction far from the minimum, switching continuously to the Gauss-Newton algorithm as the minimum is approached. This is achieved by decreasing a scalar ν, known as the Levenberg parameter, after a successful iteration, but increasing it if an uphill step is taken. The following system of equations is solved for $\Delta\mathbf{p}$ at an iteration labeled k:

$$(\mathbf{J}_k^T \mathbf{V}_{zz}^{-1} \mathbf{J}_k + \nu_k \mathbf{D}_k)\Delta\mathbf{p}_k = -\mathbf{J}_k^T \mathbf{V}_{zz}^{-1} \mathbf{r}_k \tag{26}$$

Here, $\mathbf{D}$ denotes a square matrix of order n with diagonal elements equivalent to the diagonal elements of matrix $(\mathbf{J}_k^T \mathbf{V}_{zz}^{-1} \mathbf{J}_k)$. The improved parameter set is finally calculated:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \Delta\mathbf{p}_k \tag{27}$$

The use of ITOUGH2 for both design and analysis of a well test will be illustrated in the following sections. The problem discussed below is part of the site characterization effort for a repository of low- and intermediate level nuclear wastes in Central Switzerland, and was provided by the Swiss National Cooperative for the Disposal of Radioactive Waste (Nagra).

4. Design Calculations

4.1 Problem Description and Objectives

The presence of natural gas in the host rock for a repository of radioactive wastes may greatly affect the regional flow behavior as well as the transport of radionuclides. Furthermore, gas may be generated in the repository itself, the release of which is controlled by the two-phase flow parameters of the backfill material and the formation. Small amounts of methane have been produced during the hydraulic testing of a potential host rock in the Swiss Pre-Alps. However, no firm conclusions could be reached as to whether methane existed as a free gas phase in situ, or whether the produced methane was originally entirely dissolved in the water [*Senger and Jaquet*, 1994]. Uncertainty arose from insufficiently long duration of individual test events, inaccuracy in rate and pressure measurements, and the unfavorable nature of the formation characteristics, especially its low permeability and porosity.

The purpose of our design calculations is to develop a test sequence that would be suitable for identifying gas content and two-phase hydraulic properties in a low permeability formation. The optimum test design has to meet the following objectives:

1. Analysis of the test sequence must identify whether a free gas phase is present in the formation, or whether all gas is dissolved in the pore water and merely exsolves during the pumping period.
2. The test sequence should produce data that make it possible to discriminate between alternative models, especially the choice of the characteristic curves.
3. Model-related two-phase flow parameters are to be determined with acceptably low estimation errors. The key parameters of interest are absolute permeability, formation pressure, initial gas saturation, and gas entry pressure.

4.2 Development of Alternative Test Designs

As mentioned in section 2, we do not actually solve an optimization problem by fully parameterizing the design and minimizing one of the criteria discussed above. Instead, we present a limited set of competing test designs, and evaluate their parameter estimation and model discrimination capabilities. There are many potential test designs one could examine, including single-hole and cross-hole tests, tracer tests with different partitioning of the injected chemical between the gas and liquid phase, non-isothermal injection-withdrawal tests, and combinations thereof. In this paper we focus on testing in a single, deep borehole intersecting a tight formation where the application of more sophisticated experimental layouts is unfeasible due to technical and financial restrictions.

We compare the performance of three test designs of increasing complexity by noting their capacity to discriminate among a set of alternative conceptual models. The three designs differ in the number of test events being performed. We extend a standard hydraulic testing program, referred to as sequence 1, by adding two new sequences of events (test sequences 2 and 3). The actual testing program is preceded by pretest activities (borehole history, BH), as well as a shut-in pressure recovery period (PSR) after packer inflation.

The first sequence is a standard testing procedure which consists of a series of pulse withdrawal (PW) and a constant rate pumping test (RW), followed by a shut-in recovery period (RWS). Under single-phase flow conditions, these events can be analyzed to determine the key parameters transmissivity and formation pressure head. However, if gas is present (either in dissolved form or as a free phase), conclusive results are difficult to obtain. This is mainly due to the fact that gas bubbles are formed in the borehole interval either by gas production from the formation or due to degassing of dissolved methane as borehole pressure drops below formation pressure.

The basic idea of sequence 2, which will be added to the previous sequence, is to prevent the formation of gas bubbles in the interval by performing a water injection test. Prior to testing, it is necessary to completely release the gas that was accumulated in the test interval during sequence 1. Test sequence 2 then starts with a pulse injection test (PI) to determine the test zone compressibility for the subsequent constant rate water injection test. The duration of the constant rate water injection test (RI) should be chosen such that wellbore storage effects are superseded, allowing for an accurate determination of effective liquid permeability near residual gas saturation. The RI period, however, should not last very long in order to prevent the displacement of formation gas too far away from the borehole. Test sequence 2 is terminated with a relatively long recovery period (RIS).

Finally, in test sequence 3, the water in the test interval is flushed out by gas (W2G), and a constant rate gas injection test (GRI, GRIS) is performed. The main purpose of test sequence 3 is to identify the parametric model as well as the corresponding two-phase flow parameters. If these data are available, the results of the previous two test sequences will be more conclusive as well.

Table 2 gives a summary of all test events and the definition of the three designs referred to as design A, B, and C, respectively. The last column summarizes the main objectives of each event. It will be shown, however, that all test events have to be analyzed in a joint inversion to achieve successful parameter and model identification.

Table 2. Summary of Modeled Test Events

Seq.	Event	Dur. [h]	Time [h]	Boundary Condition	Test Activity	Main Objectives and Comments
0	BH	51	0	p=4120 kPa	drilling	pressure in open borehole
	PSR	9	51	shut-in	recovery after packer inflation	obtain initial estimate of permeability and pressure
1	PW1	3	60	p=1500 kPa	pulse withdrawal	inner zone permeability
	RW	10	63	q=-0.1 kg/min	constant rate withdrawal	collect fluid samples, determine gas-water ratio
	RWS	10	73	shut-in	shut-in recovery	formation parameters
	PW2	3	83	p=3500 kPa	pulse withdrawal	observe change to PW1
2	PI	3	86	p=4500 kPa	pulse injection	release gas from interval
	RI	5	89	q=0.10 kg/min	constant rate water injection	determine permeability near residual saturation
	RIS	10	94	shut-in	shut-in recovery	determine gas content and initial formation pressure
3	W2G	2	104	p=4120 kPa	replace water in interval by gas	interval subjected to pressure in open borehole
	GRI	5	106	q=5E-3 kg/min	constant rate gas injection	determine gas entry pressure
	GRIS	20	111	shut-in	shut-in recovery	determine two-phase parameters

Design A: Sequence 0 and 1
Design B: Sequence 0, 1, and 2
Design C: Sequence 0, 1, 2, and 3

During the design stage, a number of simplifying assumptions regarding process description and model structure are made. Once data become available, these assumptions have to be reassessed based on initial data screening, diagnostic plot analysis, and inverse modeling results. We mention here some of the modeling assumptions for the design calculations and discuss the related issues.

The formation is modeled as a homogeneous, unfractured porous medium, and flow geometry is assumed to be radially symmetric. In nature, however, the presence of fractures and local heterogeneities may induce gas accumulation and preferential flow of gas towards the pumping well, potentially associated with instabilities and intermittent flow patterns as seen in the laboratory [*Glass et al.*, 1995]. This means that the actual flow geometry is much more complicated than the modeled one, and that it may not be appropriate to describe the system in terms of average quantities. We further assume that the formation gas consists of pure methane. Increased gas flow due to Klinkenberg effects is also neglected. Considering the high pressure level, this assumption may be justified despite low porosity and permeability values. We also neglect effects of hysteresis.

The phase diagnostics in TOUGH2 are performed based on a local equilibrium assumption. This means that gas comes out of solution as soon as the solubility limit is reached under the prevailing pressure and temperature conditions, and it dissolves instantaneously if the water is undersaturated with respect to methane. While this is a valid approximation for most applications of groundwater flow in porous media, the assumption may be violated if applied to two-phase flow in rough fractures [*Geller et al.*, 1995], or a borehole interval which has a relatively large volume compared to the adjacent pore space.

Simple initial and boundary conditions are prescribed. Initial formation pressure and gas saturation are assumed to be uniformly distributed. Drilling and borehole history is simulated by injecting de-aired water under a constant over pressure, and no outer boundaries are believed to be present. Finally, it is assumed that two-phase flow can be described using one of the standard characteristic curves. We arbitrarily select the van Genuchten model (Eqs. 18-20) for the generation of synthetic data which represent the true system behavior. The base case parameter set is summarized in Table 3.

Table 3. Base Case Parameter Set and Interval Specifications

Parameter	Value
log (permeability k [m^2])	-16.00
initial formation gas pressure p_i [MPa]	4.20
initial formation liquid pressure p_{li} [MPa]	4.00
borehole history pressure p_{bh} [MPa]	4.12
log (1/α [Pa])	5.51
pore size distribution index n [-]	3.00
initial gas saturation S_{gi} [-]	0.10
residual liquid saturation S_{lr} [-]	0.25
residual gas saturation S_{gr} [-]	0.01
porosity ϕ [%]	2.00
temperature T [˚C]	25.00
log (formation compressibility c_ϕ [Pa^{-1}])	-8.00
log (test zone compressibility c_{tz} [Pa^{-1}])	-8.70
Borehole and interval information:	
depth of interval midpoint [m]	-400.00
borehole radius [m]	0.08
interval length [m]	10.00
actual shut-in volume V_{bh} [m^3]	0.10

A total of 177 synthetic pressure data are generated by forward modeling. A normally-distributed random noise with a standard deviation of 50 kPa is added to all data except for the pressures during the RW period which are expected to fluctuate with a standard deviation of 100 kPa due to gas evolving in the interval. This relatively large perturbation does not only represent a potential measurement error, but also random contributions from unexplained modeling errors. We make the conservative assumption that no gas flow data will be available. Note that reliable gas rate measurements would make it easier to identify the most likely model and its parameters. The synthetic data are then matched with six different conceptual models by estimating the parameters shown in Table 4. A comparison of the performances will tell us whether the data from a specific design are able to discriminate among alternative conceptual models. The results from the correct conceptual model are then analyzed to study the parameter estimation capability of the three designs. Estimating as many parameters as feasible is a conservative approach in the sense that model discrimination becomes more difficult and the elements of the covariance matrix increase with the number of parameters. For a complete picture of conceptual model and parameter sensitivities, the procedure should be repeated using synthetic data generated with each of the alternative conceptual models. In this paper we are only discussing synthetic data produced with the van Genuchten model, and how they can be matched using the alternative model structures.

Table 4. Conceptual Models Used to Match Synthetic Data and Model Parameters Adjusted

Case	Conceptual Model	Parameters to be Estimated
VG	Two-phase van Genuchten (Eq. 18-20) #	$\log(k)$, p_i, p_{bh}, ϕ, $\log(c_{tz})$, S_{gi}, $\log(1/\alpha)$, n, S_{lr}, S_{gr}
SP	Single-phase liquid, no dissolved gas Brooks-Corey for Sequence 3	$\log(k)$, p_i, p_{bh}, $\log(c_\phi)$, $\log(c_{tz})$, $\log(p_e)$, λ, S_{lr}, S_{gr} @
DG	Single-phase liquid, degassing occurs Grant (Eq. 21, 22, 24)	$\log(k)$, p_i, p_{bh}, $\log(c_\phi)$, $\log(c_{tz})$, $\log(p_e)$, λ, S_{lr}, S_{gr} @
LI	Two-phase linear model (Eq. 13-15)	$\log(k)$, p_i, p_{bh}, ϕ, $\log(c_{tz})$, S_{gi}, $\log(p_e)$, $\log(p_{\max})$, S_{lr}, S_{gr}
BC	Two-phase Brooks-Corey (Eq. 21-23)	$\log(k)$, p_i, p_{bh}, ϕ, $\log(c_{tz})$, S_{gi}, $\log(p_e)$, λ, S_{lr}, S_{gr}
GR	Two-phase Grant (Eq. 21, 22, 24)	$\log(k)$, p_i, p_{bh}, ϕ, $\log(c_{tz})$, S_{gi}, $\log(p_e)$, λ, S_{lr}, S_{gr}

The VG model was used to generate the synthetic pressure data with the parameter set of Table 3
@ Two-phase flow parameters of model SP and DG are estimated only for Design C
Parameters are defined in Tables 1 and 3

4.3 Model Discrimination

We first examine the ability of each design to discriminate among different models. Figures 2 - 4 show the simulated interval pressures for designs A, B, and C after matching the synthetic data with each conceptual model listed in Table 4. The goodness-of-fit is measured by the estimated error variance, and the ratios s_{0i}^2 / s_{0VG}^2 are evaluated (Table 5) to see how strongly the wrong conceptual models are rejected, if at all, and whether adding test sequences 2 and 3 makes the selection of the correct van Genuchten model more stringent. If only test sequence 1 is performed, the pressure data can be matched with any of the proposed conceptual models (Figure 2). All ratios s_{0i}^2 / s_{0VG}^2 are smaller than the quantile of the F-distribution on the 95% confidence level, indicating that the estimated error variances do not significantly deviate from the one realized with the van Genuchten model. While the observation of gas flow at the surface may render the SP model unacceptable, the single-phase DG model is still a valid option, where the methane degasses during pumping. In other words it cannot be decided whether or not there is a free gas phase in the pore space under in-situ conditions.

Design B produces more selective data. The two models assuming single-phase liquid conditions, model SP and DG, are not able to match the data anymore. In addition, the two-phase model GR fails to reproduce the data (see Table 5). The high gas mobility of the GR model requires a low absolute permeability to match the strong pressure drawdown during the RW period. However, this low permeability would lead to an overprediction of the pressures during the water injection test of Design B. The simultaneous inversion of both events leads to a parameter set that balances the two counteracting behaviors, thus revealing the flaw of the model which is not seen if each event is analyzed separately. The ratio of the estimated error variances of 3.7 is larger than the quantile, suggesting that the VG performs significantly better than the GR model.

Finally, the gas injection test during sequence 3 of design C further reduces the possibility that the data are matched with a wrong conceptual model. It is interesting to note that the LI model with the simple linear characteristic curves comes closest to matching the data. However, due to the increased data base, the F-quantile becomes small, making the identification of the correct model definite.

Testing the model discrimination capability of a design as discussed here is incomplete. The number of conceptual models is theoretically infinite, and it is very likely that ambiguous solutions can be found even for design C. In our view, test design should be based on a physical understanding of the system behavior rather than on statistical measures. Proposing a sequence of water and gas injection tests is motivated by the recognition that selective data will be produced due to a decoupling of absolute permeability and total mobility. Then, the statistical measures are used to test this hypothesis, taking into account the potential uncertainty of both the data and the estimated parameters.

Table 5. Ratio of Estimated Error Variances s_{0i}^2 / s_{0VG}^2

Model	Design A	Design B	Design C
VG	1.00 +	1.00 +	1.00 +
SP	1.29 +	4.09 -	7.63 -
DG	1.16 +	4.09 -	6.38 -
LI	1.20 +	1.20 +	1.69 -
BC	1.11 +	1.34 +	2.19 -
GR	1.20 +	3.70 -	7.77 -
$F_{95\%}$	1.46	1.35	1.28

+/- Model accepted/rejected

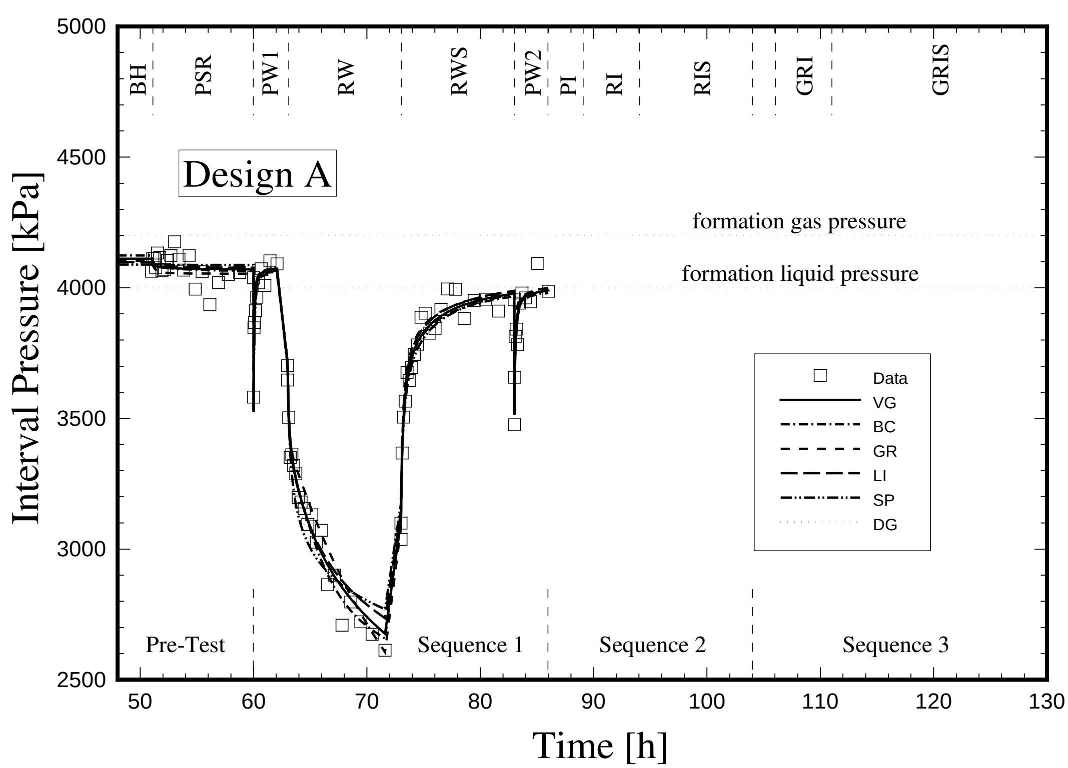

Figure 2. Matching synthetic data using six different conceptual models, Design A.

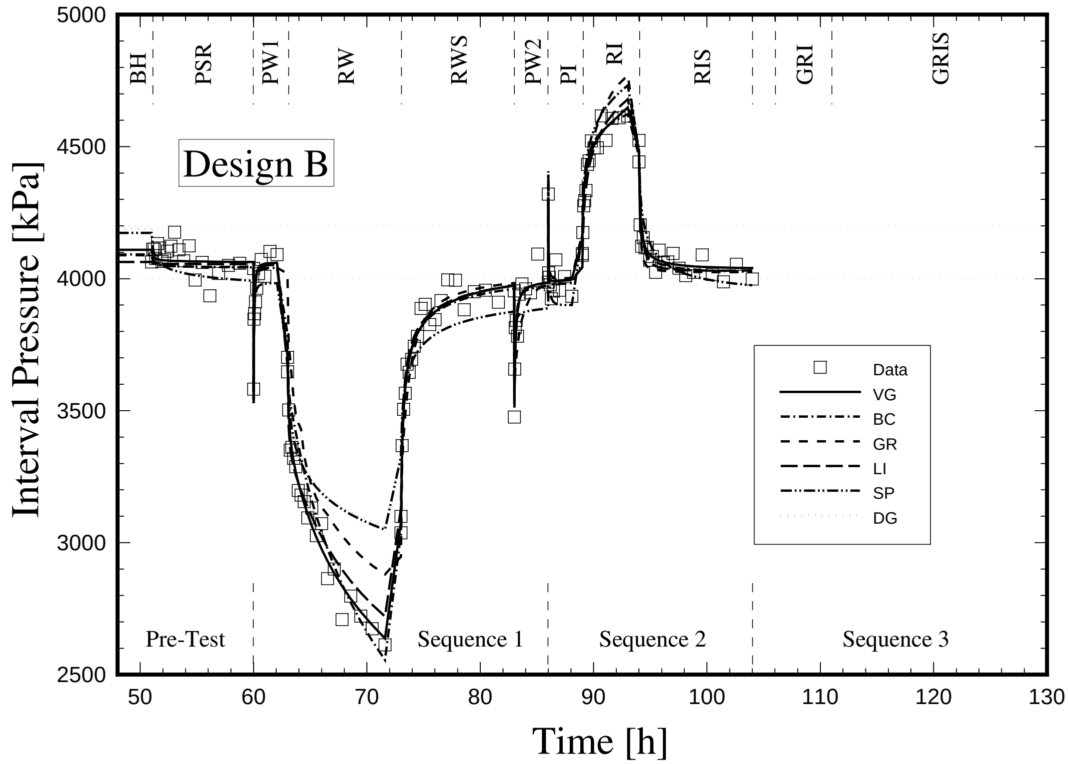

Figure 3. Matching synthetic data using six different conceptual models, Design B.

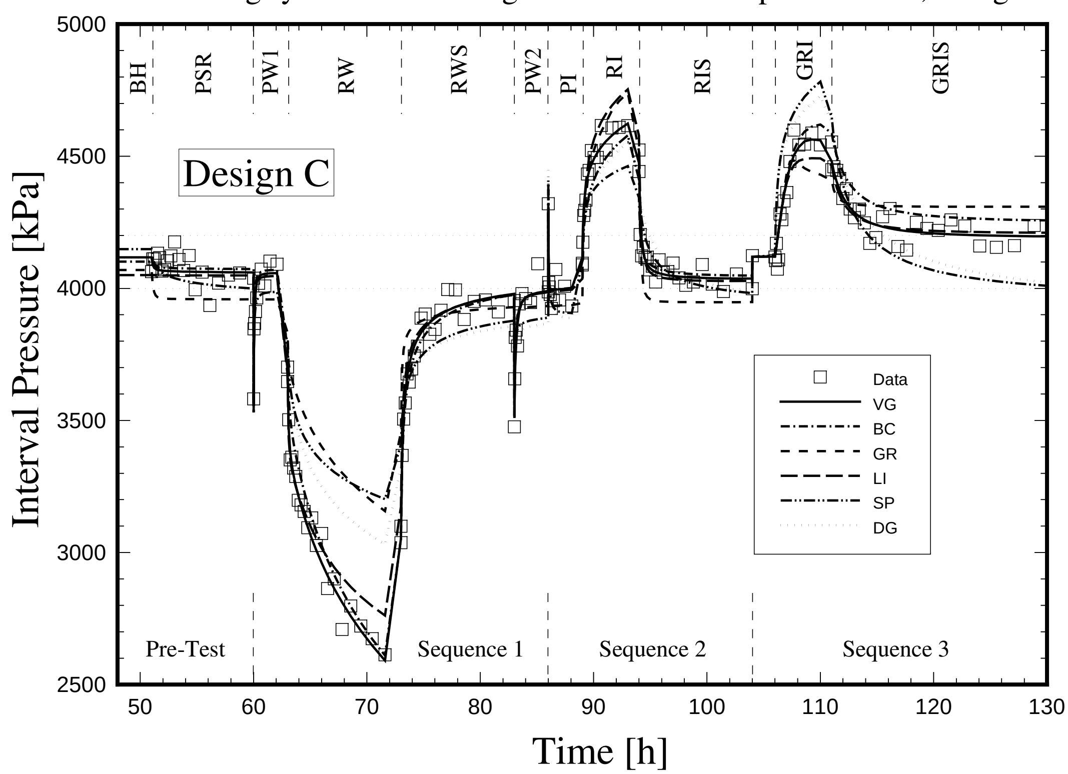

Figure 4. Matching synthetic data using six different conceptual models, Design C.

4.4 Parameter Uncertainty

The second criterion for assessing experimental designs is the parameter estimation capability which is based on an analysis of the covariance matrix of the estimated parameters. The standard deviations are summarized in Table 6. Table 7 contains the ratio χ of the marginal and joint standard deviation. A value of χ close to 1 indicates an independent estimate, whereas values close to zero suggest strong correlations of the corresponding parameter with all the other parameters subjected to the estimation process.

Table 6 demonstrates that performing test sequence 1 alone is insufficient to identify most of the key parameters, including the initial gas saturation and the formation pressure. These parameters exhibit very large estimation uncertainties which are mainly due to strong correlations as indicated by small χ-values. Parameter correlations depend on the available data. The pressure transient is affected by the total diffusivity which is increased, for example, by a decrease of system compressibility. This increase can be partly compensated by a corresponding adjustment of effective permeability. Under two-phase flow conditions, system compressibility is predominantly a function of initial gas saturation and the formation pressure which determines gas compressibility. Gas saturation also influences the effective permeability which is a combination of absolute permeability and the relative permeability function. These interferences result in strong parameter correlations, which lead to high estimation uncertainties, large eigenvalues of the covariance matrix, and a large confidence region as measured by the D-optimality criterion. Improving the test design therefore aims at reducing parameter correlations by generating data which allow for an independent estimation of the key parameters such as permeability, initial gas saturation and formation pressure. This is partially accomplished by performing a water injection test (sequence 2) which allows for a more independent estimation of absolute permeability due to the fact that saturated conditions are enforced in the vicinity of the borehole. Obtaining an accurate permeability estimate is crucial for the estimation of all the other parameters. Their respective χ-values are increased, and standard deviations below an acceptable limit are realized.

Performing a gas injection test (design C) into the almost fully saturated borehole environment further reduces the estimation uncertainty of the two-phase flow parameters. The fact that the pressures recover toward the liquid pressure at the end of the RWS and RIS periods, and toward the gas pressure at the end of the GRIS event provides a means for determining the initial gas saturation once the entry pressure is estimated. We can conclude that a simultaneous inversion of all test events draws information from data that are obtained over the entire saturation range, making possible the estimation of two-phase hydraulic properties as well as in-situ conditions.

Table 6. Standard Deviations of Estimated Parameters and Values of Optimality Criteria

Parameter	Design A	Design B	Design C
$\log(k)$	0.26	0.04	0.03
p_i	N/D	0.23	0.09
p_{bh}	0.29	0.18	0.10
S_{gi}	N/D	0.04	0.02
$\log(p_e)$	N/D	0.07	0.04
n	N/D	0.26	0.23
S_{lr}	N/D	N/D	0.18
S_{gr}	N/D	0.01	0.01
ϕ	N/D	0.01	0.01
$\log(c_{tz})$	N/D	0.13	0.13
log(D-opt.)	-24.04	-37.13	-38.76
log(A-opt.)	2.50	0.46	-0.04
log(E-opt.)	1.63	0.54	-0.67

N/D: not detectable; standard deviation larger than parameter value

Table 7. Correlation Measure χ

Parameter	Design A	Design B	Design C
$\log(k)$	0.04	0.14	0.16
p_i	0.01	0.26	0.57
p_{bh}	0.21	0.25	0.46
S_{gi}	0.03	0.16	0.18
$\log(p_e)$	0.04	0.24	0.45
n	0.01	0.25	0.27
S_{lr}	0.02	0.14	0.16
S_{gr}	0.05	0.21	0.34
ϕ	0.02	0.12	0.15
$\log(c_{tz})$	0.28	0.96	0.97

χ: ratio of standard deviation from marginal and joint probability density function.

4.5 Sensitivity Analysis

In this section we discuss some aspects of the sensitivity analysis. Figure 5 shows the scaled sensitivity coefficients (Eq. 5) of three selected parameters as a function of time. The partial derivatives $\partial z_i / \partial p_k$ were calculated numerically as a centered finite difference quotient with a perturbation of 0.005 times the respective parameter value.

The sensitivity of the observed interval pressure with respect to the undisturbed, initial formation pressure p_i shows the expected behavior with an increasing information content of the data at the end of each test event. Long term data collection following the testing program would allow for an accurate estimation of the in-situ pressure as indicated by the non-zero sensitivity coefficient at the end of the GRIS event. Unlike in single-phase liquid systems, the impact of the borehole history pressure p_{bh}, i.e. the pressure conditions in the interval during drilling and pretest activities, on the observed interval pressure is very persistent. This is a result of the fact that the history pressure controls the amount of gas being displaced during pretest activities, affecting the radius of the phase boundary. It indirectly determines the amount of gas produced during the RW period, which keeps influencing interval compressibility and thus pressure behavior throughout the experiment. Borehole history effects do not vanish until the end of the gas injection test where the phase boundary has been destroyed. The significance of the borehole history pressure for parameter estimation in a two-phase flow environment highlights the importance of accurately monitoring pretest activities.

The impact of the initial gas saturation S_{gi} on the observed interval pressures is many-fold due to its effect on overall system compressibility, gas content in the interval, relative permeability, and capillary pressure. Both the water and the gas injection tests of design B and C produce data that are sensitive with respect to changes in the initial gas saturation. Since the gas-filled interval recovers to the formation gas pressure at the end of the GRIS event as seen in Figure 4, the corresponding sensitivity coefficients remain non-zero, contributing to the estimation of the initial gas saturation. Recall, however, that sensitivity is only a necessary condition for parameter identifiability. The second requirement is independence. The sensitivity plot indicates, for example, that the two parameters history pressure and gas saturation are effectively decoupled during test sequence 2, where the sign of the sensitivity coefficient changes for gas saturation, but remains positive for the history pressure.

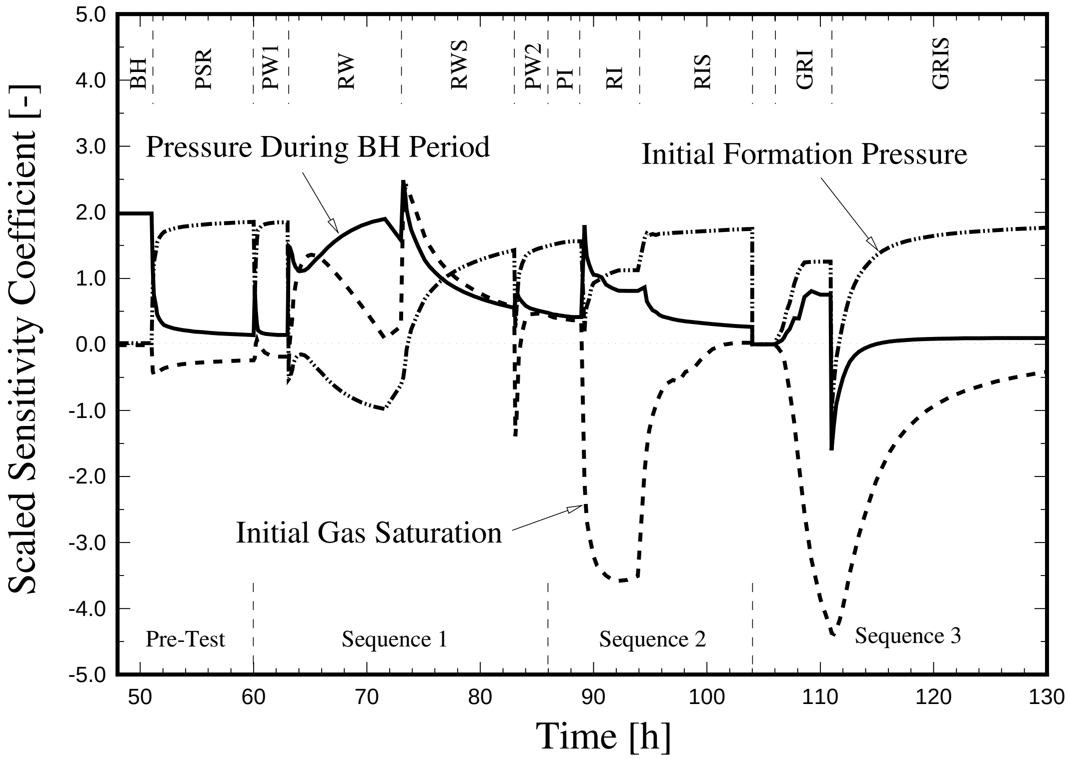

Figure 5. Scaled sensitivity coefficients for three parameters (history pressure, initial formation pressure, and initial gas saturation) as a function of time.

5. Well Test Data Analysis

5.1 Data Review

In October 1995, Nagra conducted a series of tests in a slanted borehole at a depth of about 330 m below ground surface to specifically address the issue of two-phase flow at a potential repository site for low and intermediate level radioactive wastes in Central Switzerland. The test sequence followed in part the proposal discussed in section 4. The aim is to determine whether the gas observed at the wellhead is dissolved under in-situ conditions, or present as a free phase. Furthermore, permeability, formation pressure, and two-phase flow parameters are to be estimated. We first review the information from the Quick Look Report [*QLR*, 1995] which was compiled by the field contractors. The borehole intersects the so-called Palfris formation which consists of inter-layered clayey and carbonaceous marls with calcite veins. A 2.5 m thick ductile shear zone was also identified which consists of a sequence of fractured cataclastic wallrock adjacent to thin, filled fault gouges. The appropriateness of an equivalent porous medium (EPM) approximation for such a system has been discussed by *Nagra* [1993]. If employing the simplest possible model which treats the formation as a porous medium, then the estimated values represent effective parameters under the EPM assumption, i.e., porosity and residual liquid saturation are expected to be higher than those of the intact marl matrix. Flow geometry is assumed to be radial, and a skin zone with a radius of 0.2 m is assigned to allow modeling of a region around the borehole that is disturbed due to drilling and previous test activities. The presence of a skin zone is also evident from diagnostic plots of the PSR, SWS, and RWS periods [*QLR*, 1995].

The test analyzed here is the first of several straddle packer tests performed after total depth of the borehole was reached and geophysical surveys were terminated. The interval was hydraulically tested previously, and a series of fluid loggings were performed. As a result, the borehole history is very long (about 2100 hours) and complex, consisting of 28 individual events. The initial series of test events (see Table 8) follows the proposed test sequence with a slug withdrawal test (SW, SWS) added after the first pulse test. Note that test sequence 1 leads to interval pressures that are below formation pressure. Test sequence 2 comprises a pulse injection (PI) and water injection test (RI), followed by a shut-in recovery period (RIS) and a final pulse withdrawal test (PW3). No gas injection test (sequence 3 of design) was performed.

Table 8. Test Events

Seq.	Event	Duration [h]	Time [h]	Boundary Condition	Comment
0	BH	2104.0	0.0	variable	includes drilling, hydraulic testing, fluid logging, packer inflation
1	PSR	8.9	2104.0	V_{bh}=0.1142 m^3	shut-in
	PW1	5.9	2112.9	p=2775 kPa	pulse test
	SW	9.1	2118.8	p=2169 kPa	slug withdrawal
	SWS	17.6	2127.8	shut-in	shut-in recovery
	RW	21.7	2149.8	q=prescribed	constant rate withdrawal
	RWS	20.4	2167.1	q=0.0 kg/s	shut-in recovery
	PW2	3.1	2187.6	p=2501 kPa	pulse withdrawal
2	PI	8.2	2190.7	p=3580 kPa	pulse injection
	RI	7.0	2201.0	q=prescribed	constant rate water injection
	RIS	14.5	2205.9	q=0.0 kg/s	shut-in recovery
	PW3	2.4	2220.3	p=2502 kPa	pulse withdrawal

Assessing the uncertainty of the data and detection of potential systematic errors is essential for inverse modeling. When comparing observed and calculated values, it is assumed that the two quantities are in principle compatible with each other. The numerical model produces results that represent downhole conditions, i.e. fluid flow along the borehole is not explicitly simulated. Pressure measurements were taken downhole (the sensor carrier being a few meters above the midpoint of the interval) and are considered reliable. A prior standard deviation of 50 kPa is assigned, representing the expected quality of the match. The pressure response during the RW event is unstable due to emerging gas bubbles. After an initial drawdown, pressures seem to stabilize and start to increase despite continuous production. At this point in time, the pumping rate was tripled, leading to a similar behavior at lower pressures. Gas and liquid flow rates are measured at the land surface near atmospheric conditions. Presumably there is a time lag between the occurrence of gas in the test interval and its manifestation at the surface. Furthermore, the fluid mixture was received in discontinuous slugs at the flow meters. Instead of calibrating the model against these highly fluctuating flow rate measurements, a moving average was taken and integrated to obtain the cumulatively produced volume of gas under normal conditions. A standard deviation of 5.0 liters was assigned to these derived data points. In summary, the data used for model calibration are the interval pressure during the entire test period, and the cumulative gas production during the RW event.

5.2 Inverse Modeling Results

A number of inversions were performed in an attempt to match the borehole data using different model structures and assumptions regarding the origin of the observed gas. The model assumptions are the same as described previously (see Table 4). We first examine the overall goodness-of-fit (Table 9), expressed by the estimated error variance s_0^2 and the mean square residual of pressure and cumulative gas production. Recall the respective prior standard deviations of 50 kPa and 5 normal liters; the prior error variance σ_0^2 is equal to 1.0 by definition. From Table 9 we conclude that the two-phase Brooks-Corey (BC) model performs best. However, with a mean square residual of 78.8 kPa and 8.9 liters, it did not quite achieve the expected match, as is evident from a s_0^2-value which is significantly greater than 1.0. This may indicate that the assumptions regarding measurement errors were too optimistic. The systematic deviations shown in Figures 6 and 7 below suggest, however, that the simplifications in the conceptual model inhibit accurate reproduction of all aspects of the system behavior. The model assumptions and their potential impact have been discussed in sections 4.2 and 5.1, and will be further scrutinized below for the selected model.

Figures 6 and 7 show the comparison between observed and simulated gas production and interval pressures, respectively. It is obvious from Figure 6 that degassing of dissolved methane does not produce the observed amount of gas despite the assumed high gas relative mobility, whereas an almost perfect match was obtained with both the Brooks-Corey and van Genuchten model assuming a free gas phase. Figure 6 also shows the measured flow rates to illustrate the noisy character of the data. Calibrating against the cumulative gas production instead of the flow rates makes the inversion more stable.

Matching pressures over the entire test duration is very difficult and requires making a number of assumptions regarding the fate of the gas in the interval. Note that the recovery pressure seems to decrease with each additional event of test sequence 1. This could be interpreted as a boundary effect where a zone of reduced permeability is encountered. However, pressure recovers very fast to a higher level during the RI and RIS periods, suggesting that another process such as gas entrapment in the interval is responsible for the apparent boundary effect. The data from test sequence 2 prove to be selective regarding assumptions about the conceptual model.

Table 9 and Figure 7 below reveal the interesting fact that the degassing model DG and the single-phase model SP (not shown in Figures 6 and 7) provide a match to the pressure data that is as good as the one obtained with the two-phase models; the single-phase models evidently fail to reproduce the gas flow data. The favorable pressure matches of SP and DG were only achieved, however, by allowing for an unreasonably high pore space compressibility of about 10^{-7} Pa^{-1}. This high system compressibility can be taken as an indication that there is actually a free gas phase present in the formation which needs to be accounted for by high pore space compressibilities in the two models that do not explicitly consider free gas. Similarly, a high and time-varying system compliance of the borehole is required for the single-phase models to account for entrapped gas in the interval, whereas a lower and constant value was obtained with the two-phase model. Note that the simple linear model (LI, not shown in Figures 6 and 7) performs second best after the Brooks-Corey (BC) model, in agreement with the results from the design calculations. While the gas production data are matched reasonably well with the van Genuchten (VG) model, it fails to reproduce the pressure data mainly due to its lack of a finite gas entry pressure. Gas is more mobile for weak suction pressures, i.e. the required amount of gas is produced with less pressure drawdown, and generally higher recovery pressures are achieved. Another interesting result was obtained with the Grant (GR) model (not shown in Figures 6 and 7) which tends to overpredict gas flow due to its high gas relative permeability. In order to match cumulative gas production, a very low permeability value is needed which in turn leads to a contradiction with the observed pressure transient during the RI period. As a result, the overall performance is very poor, demonstrating the selective character of a joint inversion of all test events, especially test sequence 2. It should be mentioned in this context that better fits can be obtained when using data from test sequence 1 only. As discussed in section 4, these inversions are less conclusive. Additional selective information is contained in the gas flow data which help discriminate between the alternative scenarios.

Table 9. Ratio of Estimated Error Variances and Mean Square Residual

		Root Mean Square Residual	
Model	s_0^2	Pressure [kPa]	Gas Flow [l]
SP	27.2	70.7	186.1
DG	26.2	75.0	177.6
LI	3.3	90.5	12.4
VG	25.2	342.5	19.5
BC	2.7	78.8	8.9
GR	51.8	574.2	34.7
a priori	1.0	50.0	5.0

Discarding the degassing scenario and having identified the Brooks-Corey characteristic curves as the most likely model, we can finally discuss the estimated parameter set shown in Table 10. The estimates for the skin zone indicate higher permeability and lower capillary strength than the surrounding rock. This leads to an accumulation of gas near the borehole, affecting the early-time behavior of all test events conducted after the RW period. A relatively low but nevertheless significant initial gas saturation of 9 % was determined. The estimated gas entry pressure of 0.28 MPa is lower than the values obtained in an undisturbed section of the same formation [*Adams and Wyss*, 1994], as expected. A more accurate estimate of this parameter would be possible with data from a gas injection test. Initial gas and liquid pressures are 3.72 and 3.42 MPa, respectively. The high residual water saturation of 0.46 is justified by the EPM assumption. Note, however, that only data in the high saturation range are available which is also one of the reasons for the poor identification of the pore size distribution index λ. Finally, the estimate of the system compliance is consistent with its measured value [*QLR*, 1995].

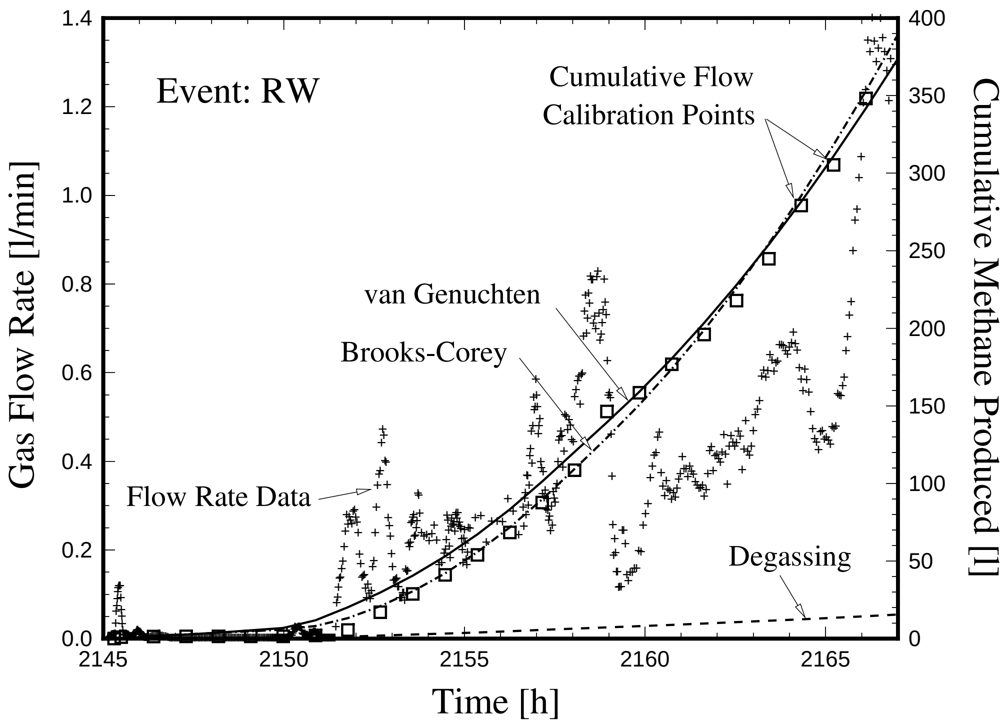

Figure 6. Comparison between observed and calculated cumulative gas produced during the RW pumping period, model BC, VG and DG. Data supplied by Nagra.

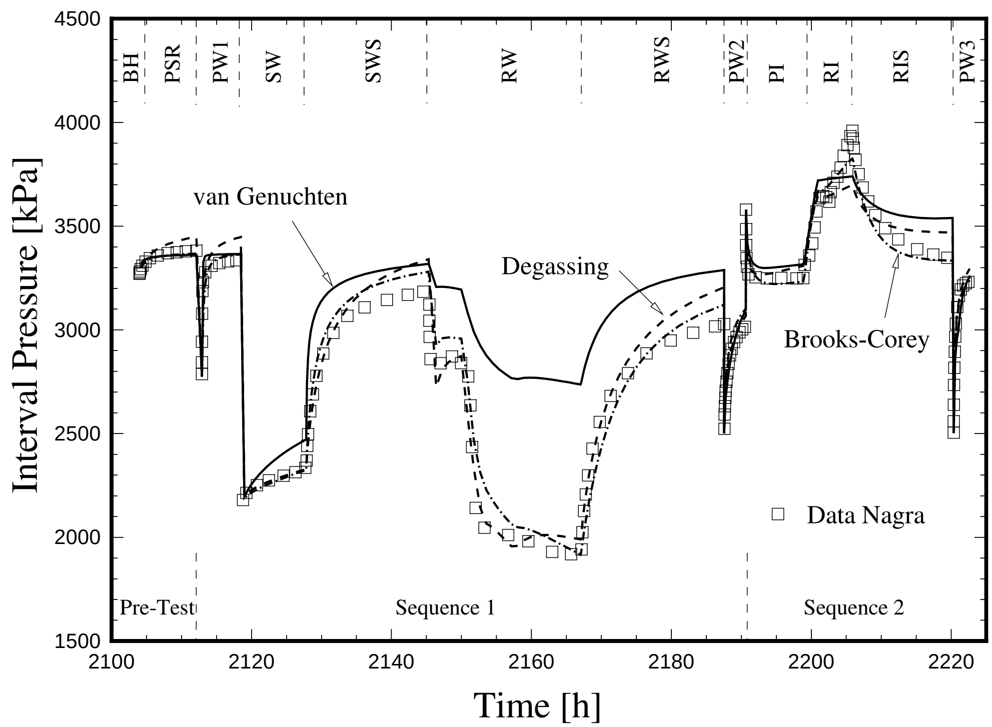

Figure 7. Comparison between observed and calculated pressures, model BC, VG and DG. Data supplied by Nagra.

The standard deviations of the estimated parameters are very low. All parameters, with the exception of system compliance, are highly dependent on each other as indicated by the small χ-values. Recall that the standard deviation only measures a statistical uncertainty. Systematic errors, which are likely to be present as the residual pattern suggests, may lead to errors that exceed the range of the parameter uncertainty indicated here. The EPM assumption, the simplified flow geometry, wellbore conditions during and after gas accumulation, hysteresis effects, and considerable uncertainty regarding wellbore history are among the issues that need to be reconsidered. Nevertheless, the joint inversion of different data from a multitude of test events allowed us to assess the conceptual model and to identify its key parameters.

Table 10. Estimated Parameter Set

Parameter	Initial Guess	Best Estimate	Std. Dev.	χ
log($k_{formation}$ [m^2])	-16.00	-16.18	0.03	0.25
log(k_{skin} [m^2])	-16.00	-15.78	0.11	0.07
porosity ϕ [%]	2.00	3.50	0.56	0.33
S_{lr}	0.30	0.46	0.09	0.02
S_{gr}	0.05	0.03	0.01	0.03
log($p_{e,formation}$ [Pa])	5.00	5.44	0.07	0.11
log($p_{e,skin}$ [Pa])	5.00	3.19	0.13	0.05
PSDI λ	2.00	2.46	0.32	0.08
initial gas saturation S_{gi}	0.10	0.09	0.01	0.03
initial gas pressure [MPa]	3.65	3.72	0.04	0.13
log (system compliance [Pa^{-1}])	-8.00	-7.76	0.09	0.70

χ: ratio of standard deviation from marginal and joint probability density function.

6. Summary and Conclusions

It was recognized from previous studies [*Finsterle*, 1994] that performing a standard test sequence consisting of a series of pulse withdrawal and pumping tests (sequence 1) leads to indeterminate results regarding the conceptual model and two-phase parameters of a tight formation. Our design calculations using synthetic, noisy test data showed that some of the problems can be overcome by adding a second test sequence where water is injected followed by an extended recovery period (sequence 2). In order to further reduce parameter correlations and to actually determine two-phase hydraulic properties, a third test sequence was proposed which includes a gas injection test (sequence 3). Furthermore, it was demonstrated that:

1. It is very unlikely that the two-phase characteristic curves and their parameters can be identified if only data from sequence 1 are analyzed. The ability to identify the correct model increases substantially if test sequences 2 and 3 are added.

2. Reduction of parameter correlations and thus increase in estimation accuracy may allow identification of initial gas content along with permeability and formation pressure.

3. A gas injection test is required to determine additional two-phase flow parameters.

Data from a test that partially followed the proposed test sequence were analyzed. Test sequence 3 proposed in the design stage was omitted. The situation was further complicated by the extremely long borehole history period. On the other hand, the added slug test and the fact that relatively accurate gas production data were available have improved the possibility of identifying the parameters of interest. Good matches were obtained to the observed interval pressures and cumulative gas production. The following conclusions were reached:

4. Performing a water injection test was crucial to discriminate among the conceptual alternatives. There are strong indications that the gas seen at the surface does not come out of solution during pumping, but is present as a free phase under in-situ conditions.

5. Exact knowledge of borehole history and test activities affecting interval conditions is required for the analysis of pressure and gas flow data from a deep borehole.

6. The combination of sophisticated process simulation with efficient and robust inversion techniques provides a tool for the design and analysis of hydraulic test under two-phase flow conditions.

Acknowledgments. This work was supported by a grant from the Swiss National Cooperative for the Disposal of Radioactive Waste (Nagra), Wettingen, Switzerland, through U.S. Department of Energy Contract No. DE-AC03-76SF00098. Thanks are due to P. Marshall and S. Vomvoris (Nagra) for their support and for providing the field data. The careful reviews by C. Doughty and K. Karasaki (LBNL) are gratefully acknowledged.

References

Adams, J., and E. Wyss, Hydraulic packer testing in the Wellenberg boreholes SB1 and SB2, Methods and field results, *Nagra Technical Report* NTB 93-38, Nagra Wettingen, Switzerland, 1994.

Carrera, J., Estimation of aquifer parameters under transient and steady-state conditions, Ph.D. Dissertation, Dept. of Hydrology and Water Resour., University of Arizona, Tucson, 1984.

Carrera, J., and S.P. Neuman, Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information, *Water Resour. Res.*, *22*(2), 199-210, 1986.

Carrera, J., State of the art of the inverse problem applied to the flow and solute transport equations, in Groundwater Flow and Quality Modeling, edited by E. Custodio, A. Gurgui, and J. P. Lobo Ferreira, *NATO ASI Ser. C*, *224*, 549-583, 1987.

Casman, E. A., D. Q. Naiman, and C. E. Chamberlin, Confronting the ironies of optimal design: Nonoptimal sampling designs with desirable properties, *Water Resour. Res.*, *24*(3), 409-415, 1988.

Cooley, R. L., L. F. Konikow, and R. L. Naff, Nonlinear-regression groundwater flow modeling of a deep regional aquifer system, *Water Resour. Res.*, *22*(13), 1759-1778, 1986.

Dettinger, M. D., and J. L. Wilson, First order analysis of uncertainty in numerical models of groundwater flow, Part 1. Mathematical development, *Water Resour. Res.*, *17*(1), 149-161, 1981.

Finsterle, S., ITOUGH2 user's guide, Version 2.2, Report No. LBL-35581, Lawrence Berkeley National Laboratory, Berkeley, Calif., August 1993.

Finsterle, S., Inverse modeling of Test SB4-VM2/216.7 at Wellenberg, Report No. LBL-35454, Lawrence Berkeley National Laboratory, Berkeley, Calif., March 1994.

Finsterle, S., Design of a welltest for determining two-phase hydraulic properties, Report No. LBL-37448, Lawrence Berkeley National Laboratory, Berkeley, Calif., January 1995.

Finsterle, S. , and K. Pruess, Solving the estimation-identification problem in two-phase flow modeling, *Water Resour. Res.*, 31(4), 913-924, 1995.

Finsterle, S., and P. Persoff, Determining permeability of tight rock samples using inverse modeling, Report No. LBL-39296, Lawrence Berkeley National Laboratory, Berkeley, Calif., submitted to *Water Resour. Res.*, August 1996.

Geller, J. T., C. Doughty, J. C. S. Long, and R. J. Glass, Disturbed zone effects: Two phase flow in regionally water-saturated fractured rock, FY94 annual report, Report No. LBL-36848, Lawrence Berkeley National Laboratory, Berkeley, Calif., January 1995.

Glass, R. J., M. J. Nicholl, and V. C. Tidwell, Challenging models for flow in unsaturated, fractured rock through exploration of small scale processes, *Geophysical Research Letters*, *22*(11), 1457-1460, 1995.

Hamilton, D. C., and D. G. Watts, A quadratic design criterion for precise estimation in nonlinear regression models, *Technometrics*, *27*(3), 241-250, 1985.

International Formulation Committee, A formulation of the thermodynamic properties of ordinary water substance, IFC Secretariat, Düsseldorf, Germany, 1967.

James, A. L., and C. M. Oldenburg, Linear and Monte Carlo error analysis for subsurface contaminant transport simulation, Report No. LBL-38507, Lawrence Berkeley National Laboratory, Berkeley, Calif., submitted to *Water Resour. Res.*, June 1996.

Kashyap, R. L., Optimal choice of AR and MA parts in autoregressive moving average models, *IEEE Trans. Pattern Anal. Mach. Intel.*, *PAMI-4*(2), 99-104, 1982.

Kool, J. B., J. C. Parker, and M. Th. van Genuchten, Parameter estimation for unsaturated flow and transport models - A review, *J. Hydrol.*, *91*, 255-293, 1987.

Knopman, D. S., and C. I. Voss, Behavior of sensitivities in the advection-dispersion equation: Implications for parameter estimation and sampling design, *Water Resour. Res.*, *24*(2), 225-238, 1988.

Knopman, D. S., and C. I. Voss, Multiobjective sampling design for parameter estimation and model discrimination in groundwater solute transport, *Water Resour. Res.*, *25*(10), 2245-2258, 1989.

Knopman, D. S., C. I. Voss, and S. P. Garabedian, Sampling design for groundwater solute transport: Tests of methods and analysis of Cape Cod tracer test data, *Water Resour. Res.*, *27*(5), 925-949, 1991.

Levenberg, K., A method for the solution of certain nonlinear problems in least squares, *Q. Appl. Math.*, *2*, 164-168, 1944.

Luckner, L., M. Th. van Genuchten, and D. Nielsen, A consistent set of parametric models for the two-phase flow of immiscible fluids in the subsurface, Water *Resour. Res.*, *25*(10), 2187-2193, 1989.

Marquardt, D. W., An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math.*, *11*(2), 431-441, 1963.

McLaughlin, D., and L. R. Townley, A reassessment of the groundwater inverse problem, *Water Resour. Res.*, 32(5), 1131-1161, 1996.

Morgan, M. G., and M. Henrion, Uncertainty, a guide to dealing with uncertainty in qualitative risk and policy analysis, Cambridge University Press, 1990.

Nagra, Untersuchung zur Standorteignung im Hinblick auf die Endlagerung schwach- und mittelaktiver Abfälle: Geologische Grundlagen und Datensatz zur Beurteilung der Langzeitsicherheit des Endlagers für schwach- und mittelaktive Abfälle am Standort Wellenberg (Gemeinde Wolfenschiessen, NM), *Nagra Techn. Ber.* NTB 93-26, Nagra, Wettingen, Switzerland, 1993.

Pruess, K., TOUGH User's Guide, Nuclear Regulatory Commission Report NUREG/CR-4645, NRC, Washington, DC, 1987.

Pruess, K., TOUGH2 - A general-purpose numerical simulator for multiphase fluid and heat flow, Report No. LBL-29400, Lawrence Berkeley National Laboratory, Berkeley, Calif., May 1991.

Pruess, K., and T. N. Narasimhan, A practical method for modeling fluid and heat flow in fractured porous media, *Soc. Pet. Eng. J.*, *25*(1), 14-26, 1985.

QLR, Quick Look Report SB4a/s - Test VM11, Golder Associates, 1995.

Senger, R.K., and O. Jaquet, Evaluation of two-phase flow parameters for the Valanginian Marl, Nagra Internal Report, Nagra, Wettingen, Switzerland, April 1994.

Steinberg, D. M., and W. G. Hunter, Experimental design: Review and comment, *Technometrics*, *28*, 71-97, 1984.

Sun, N., and W. Yeh, Coupled inverse problems in groundwater modeling, 2. Identifiability and experimental design, *Water Resour. Res.*, *26*(2), 2527-2540, 1990.

Usunoff, E., J. Carrera, and S. F. Mousavi, An approach to the design of experiments for discriminating among alternative conceptual models, *Adv. Wat. Resour. 15*, 199-214, 1992.

Wagner, B. J., and S. M. Gorelick, Optimal groundwater management under parameter uncertainty, *Water Resour. Res.*, *23*(7), 1162-1174, 1987.

Woldt, W., I. Bogardi, W. E. Kelly, and A. Bardossy, Evaluation of uncertainties in a three-dimensional groundwater contamination plume, *J. Contam. Hydrol.*, *9*, 271-288, 1992.

Yeh, W. W.-G., Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resour. Res.*, *22*(2), 95-108, 1986.

Yeh, W. W.-G., and Y. S. Yoon, Aquifer parameter identification with optimum dimension in parameterization, *Water Resour. Res.*, *17*(3), 664-672, 1981.