Title
Design of a Welltest for Determining Two-Phase Hydraulic Properties

Author
Finsterle, S.

Publication Date
1995

LBL-37448
UC-400

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

EARTH SCIENCES DIVISION

Design of a Welltest for Determining Two-Phase Hydraulic Properties

S. Finsterle

January 1995

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

Available to DOE and DOE Contractors
from the Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (615) 576-8401

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, VA 22161

DISCLAIMER

LBL-37448
UC-400

Design of a Welltest for Determining Two-Phase Hydraulic Properties

S. Finsterle

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

January 1995

This work was carried out under U.S. Department of Energy Contract No. DE-AC03-76SF00098 for the Director, Office of Civilian Radioactive Waste Management, Office of External Relations, and was administered by the Nevada Operations Office, U.S. Department of Energy, in cooperation with the Swiss National Cooperative for the Disposal of Radioactive Waste (Nagra).

SUMMARY

This report describes the design of a well test to determine two-phase hydraulic properties of a low permeability, low porosity formation. Estimation of gas-related parameters in such formations is difficult using standard pumping tests mainly because of the strong fluctuations in the pressure and flow rate data which are a consequence of gas bubbles evolving in the test interval. Even more important is the fact that the data do not allow distinguishing among alternative conceptual models. The estimated parameters are therefore uncertain, highly correlated, and ambiguous. In this study we examine a test sequence that could be appended to a standard hydraulic testing program. It is shown that performing a series of water and gas injection tests significantly reduces parameter correlations, thus decreasing the estimation error. Moreover, the extended test sequence makes possible the identification of the model that describes relative permeabilities and capillary pressures. This requires, however, that data of high accuracy are collected under controlled test conditions. The purpose of this report is to describe the modeling approach, assumptions and limitations of the procedure, and to provide practical recommendations for future testing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Problem Description

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 properties of nuclides. 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.

Gas has been produced from the Valanginian marl at the Oberbauenstock (OBS) and the Wellenberg (WLB) site in Central Switzerland. The tests conducted in the Valanginian and Tertiary marl intervals of the boreholes at Wellenberg were reviewed and categorized to identify zones that potentially contain a free gas under natural in-situ conditions [*Lavanchy and Johns*, 1994]. A total of 17 tests were reported by the field contractors to have produced gas. However, no reliable gas and/or liquid flow data were provided for 10 tests, and 4 tests showed gas-water ratios which were lower than the saturated solution gas-water ratio, indicating that the gas probably came out of solution during pumping. Significant amounts of gas were only encountered during one test at borehole SB2 and two tests at borehole SB4. Several attempts were made to analyze test SB4-VM2/216.7 by means of numerical simulations [*Senger and Jaquet*, 1994; *Finsterle*, 1994]. These studies showed that it is very difficult to derive conclusive results regarding the gas content of the formation. Consequently, no reliable estimates of two-phase hydraulic properties of the marl could be obtained.

According to the criteria used by *Lavancy and Johns* [1994], the main reasons for weak parameter identification are insufficiently long duration of individual test events, uncertainty associated with the rate and pressure measurements, and the unfavorable formation characteristics itself, especially its low permeability and the presence of a low pressure zone. *Lavancy and Johns* [1994] have presented a number of recommendations to address these difficulties in future tests at Wellenberg.

Finsterle [1994] has shown that the observed gas-water ratio is a poor indicator for the origin of the gas. If gas is highly mobile at low saturations, relatively high gas fluxes are observed during the initial stages of the pump test, leading to gas-water ratios which are above the critical saturation, even though gas originated from degassing only. On the other hand, free gas may be trapped and immobilized, leading to gas-water ratios which are close to or below the critical value. The solution thus depends on the characteristic model and the flow behavior near residual gas saturation which is highly uncertain. Extended pumping periods with nearly-steady production rates are required to conclusively determine the origin of the gas from gas-water ratio measurements. *Finsterle* [1994] drafted a revised test sequence which is likely to reduce the ambiguity in the gas flow data and which improves the ability to identify two-phase flow parameters.

The difficulty of determining gas saturations at Wellenberg basically stems from the considerable uncertainties associated with the data, the parameters, and the assumptions of the model used to analyze the data. Therefore, a test design has to be conceived which:

(1) reduces the risk of producing unreliable and noisy data,

(2) enables independent determination of model parameters, and

(3) is not critically dependent upon potentially controversial model assumptions, or which distinguishes conclusively between competing alternatives.

1.2 Objectives

The purpose of this study is to design a test sequence that is 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 provide a criterion that allows one to distinguish between a free gas content in the formation, and gas that is originally dissolved in the pore water, but comes out of solution during the pumping period.

(2) The test sequence should produce data that make possible the discrimination between alternative models, i.e. the data have to be sensitive to the choice of the characteristic curves (e.g. van Genuchten vs. Brooks-Corey) in order to be selective.

(3) Model-related two-phase flow parameters are to be determined, i.e. estimation errors have to be acceptably low. Furthermore, the estimates should be as independent as possible.

(4) The test sequence should be designed as an extension of the standard test procedure. Criteria have to be developed to decide whether to invoke extended testing.

(5) The test sequence should be technically feasible, simple, and economical.

In order to meet these objectives, a procedure has to be conceived, furnishing suitable performance measures, based on which the test design can be evaluated. The methodology will be discussed in Section 3, using the test sequence described in Section 2. A reference case is developed in Section 4, and the selection of an improved test design is discussed in Section 5, followed by conclusions and recommendations for future testing (Section 6).

2 PROPOSED TEST SEQUENCE

Test design is an iterative optimization process which comprises problems of hydrological, technical, and economic nature. The optimization problem is thus difficult to solve because some of the objectives cannot easily be quantified. Furthermore, weighting goals of different nature are based on a number of management decisions that cannot be incorporated into an optimization model, yet they strongly affect the final design.

The test design discussed herein is a compromise. It was recognized that optimizing the test sequence with respect to parameter identification alone leads to a solution that is technically demanding (if not unfeasible) and financially unacceptable. It was therefore decided to approach the problem in three stages as visualized in Figure 1 (acronym test names are defined in the glossary).

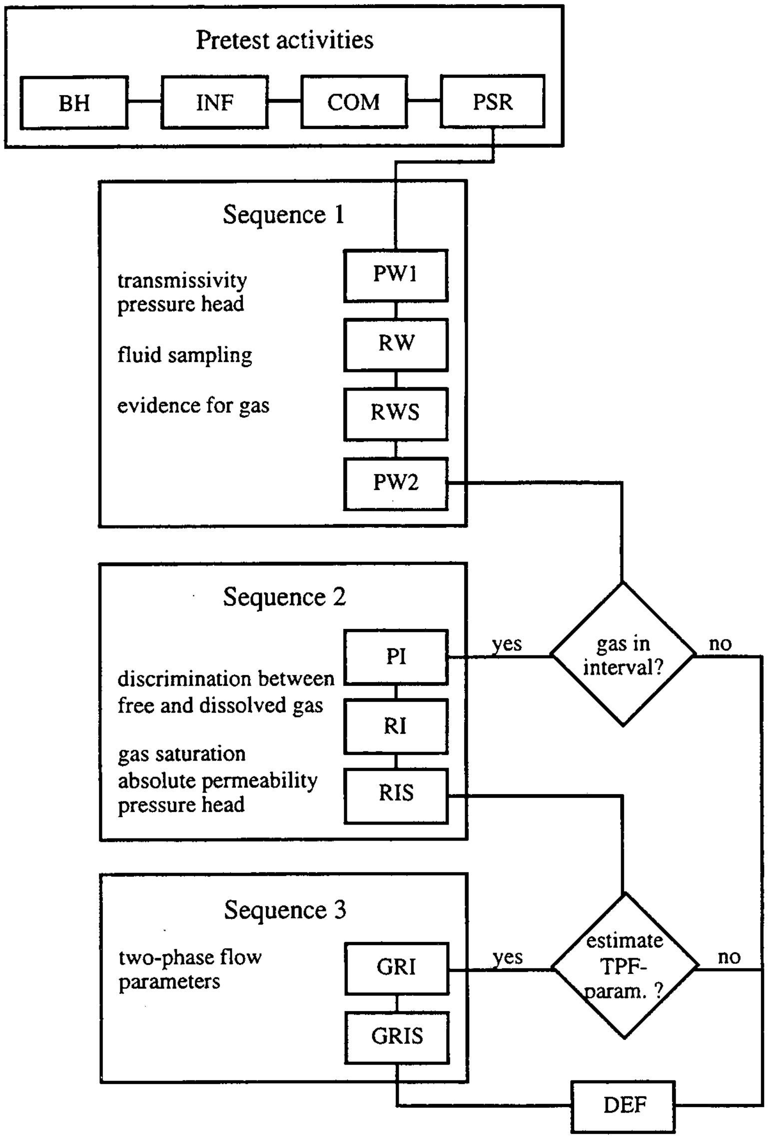

Figure 1: Proposed test sequence and major objectives

In the first sequence following pretest activities (BH, INF, COM, PSR), a standard testing procedure is conducted which consists of a series of pulse withdrawal (PW) and constant rate pumping tests (RW, 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), it is obvious from the study conducted by *Lavanchy and Johns* [1994] that 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 at least due to degassing of dissolved methane caused by the imposed pressure reduction. This usually induces noisy readings and thus unreliable pressure and flow rate data [*Adams and Wyss*, 1994]. Furthermore, the gas in the interval is a complicating factor for analysis and may lead to erroneous results if not explicitly accounted for. The presence of gas may be directly observed at the surface or detected during the PW2 event which will show a significantly increased total system compressibility if a free gas phase appears in the test interval.

Test sequence 2 can be invoked if either of the three following criteria is met:

(1) gas is produced during the RW period,

(2) diagnostic analysis of the RWS event shows a composite system behavior which is likely to be caused by a phase boundary,

(3) the diagnostic plots of either of the events of test sequence 1 indicate atypical behavior, i.e. cannot be analyzed using standard interpretation techniques such as type curve matching or the use of a single-phase wellbore simulation program.

The basic idea of sequence 2 is to prevent degassing and 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 make sure that no gas stayed in the interval, and 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 an accurate determination of liquid permeability. The RI period, however, should not last very long in order to prevent the displacement of formation gas too far away from the borehole. The subsequent, relatively long recovery period (RIS) will show a composite system behavior if free gas is present. The possibility of determining the actual gas content is increased due to the fact that more reliable estimates of permeability and pressure head are obtained. This is important because the gas content estimate is strongly correlated to these two parameters.

Finally, a gas injection test (GRI, GRIS) can be performed (test sequence 3). The main purpose 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. Performing a gas injection test requires that the water in the interval be displaced by gas using a control line; the procedure is described in *Enachescu et al.* [1992].

Table 1 gives a summary of all test events considered in the simulation. The duration of each event depends on the actual formation characteristics, especially on the permeability. The test sequence presented in Table 1 refers to the base case parameter set discussed in Section 4.

Table 1: Summary of modeled test events

Sequ.	Event	duration [h]	Time [h]	Boundary Condition	Comment
0	BH	48	0	p=4120 kPa	pressure based on depth=400 m and drilling fluid density ρ=1050 kg/m^3
	INF COM1	3	48	p=4120 kPa	inflate packers, compliance period
	PSR	9	51	shut-in V_{bh}=0.1 m^3	obtain initial permeability and pressure head estimate
1	PW1	3	60	p=1500 kPa	determine test zone compressibility and obtain permeability estimate
	RW	10	63	q=-0.1 kg/min	measure gas-water ratio; obtain permeability and gas content estimates; collect fluid samples
	RWS	10	73	shut-in	obtain permeability, gas content, and pressure head estimates
	PW2	3	83	p=3500 kPa	determine whether free gas is in the test interval
2	PI	3	86	p=4500 kPa	release gas from test interval and determine test zone compressibility
	RI	5	89	q=0.10 kg/min	determine absolute permeability
	RIS	10	94	shut-in	determine gas content and pressure head
3	COM2	2	104	p=4120 kPa	displace water in interval by gas
	GRI	5	106	q=4.5E-3 kg/min	identify parametric model; determine threshold pressure and two-phase flow parameters
	GRIS	20	111	shut-in	determine two-phase flow parameters
	DEF	0	131		deflate packers

3 MODEL ASSUMPTIONS

3.1 General Methodology

The appropriateness of the test design with respect to the objectives mentioned in Section 1.2 is examined by means of numerical modeling in combination with linear sensitivity and error analysis. We used the TOUGH2 code [*Pruess*, 1991] to simulate two-phase flow phenomena, including phase interferences, capillary forces, and the dissolution and degassing of methane, the main constituent of the gas encountered at Wellenberg. The calculation of sensitivity and correlation coefficients is an option provided by the ITOUGH2 code [*Finsterle*, 1993; *Finsterle and Pruess*, 1995].

The general methodology used in this study is to model the proposed test sequence in a forward mode, thus generating synthetic data as they might be observed in the borehole. Subsequently, this data is used to solve the inverse problem using ITOUGH2. Only one iteration is required since the true parameter combination is already known. The *a posteriori* error analysis provides the estimation error of the parameters as well as their correlation structure. A change in the test design will result in a different covariance matrix for the parameters of interest. The optimum test design is the one that yields the lowest parameter uncertainty. This is usually achieved by producing more data that are sensitive with respect to the parameters of interest. At the same time, the data should allow for a more independent estimate of the model parameters, i.e. the correlations among the parameters is reduced. Independent parameter estimation means that the *joint* standard deviation (the standard deviation of a parameter taking into account the uncertainty of all the other parameters) approaches the *conditional* standard deviation (the standard deviation of a parameter assuming that all the other parameters are exactly known). Generally, neglecting the influence of uncertain parameters that are correlated to the parameters of interest leads to overly optimistic conclusions regarding the accuracy with which parameters can be determined. The approach presented here allows for a more objective evaluation of alternative test designs.

Both the modeling of two-phase flow in the vicinity of a borehole as well as the subsequent error analysis are based on a number of assumptions which will be critically reviewed in the remainder of this section. A discussion of the proposed approach, definitions of the statistical measures as well as their interpretation can be found in Appendix A.

3.2 Modeling of Tests Under Two-Phase Flow Conditions

A technical description of the TOUGH2 code, including a discussion of the physical processes modeled as well as the mathematical and numerical methods is provided in the TOUGH and TOUGH2 user's guides [*Pruess*, 1987; 1991]. An equation-of-state module that describes the thermophysical properties of methane-water mixtures is linked to the program. Various simplifying assumptions are made which potentially influence the simulation results. They are:

(1) 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 a borehole interval which has a relatively large volume compared to the adjacent pore space. Individual gas bubbles entering the borehole interval may not dissolve instantaneously because of the limited interface area between gas and liquid.

(2) We will model the formation as a homogeneous, unfractured porous medium, and flow geometry is assumed to be radially symmetric. In nature, however, local heterogeneities may induce gas accumulation and preferential flow of gas towards the pumping well, associated with instabilities and intermittent flow patterns. These phenomena not only lead to noisy field data but also to a system behavior which cannot be described in terms of average quantities.

(3) In the model, the state variables represent effective values under downhole conditions. They usually are not directly comparable to the pressures and flow rates observed at the surface. Degassing of uprising water, phase segregation and other effects such as air-lifting of water in the tubing have to be accounted for when comparing observed data and model results. This requires additional assumptions which may induce further uncertainties or systematic errors.

(4) We will make a number of assumptions regarding the borehole conditions (e.g. free gas can be completely released between certain test events; test zone compressibility is small and constant with time, etc.). Simple initial and boundary conditions will be prescribed (drilling and borehole history can be modeled as injection of de-aired water; formation pressure and gas saturation is initially uniformly distributed; no outer boundaries are present, etc.). During pumping, the composition of the produced fluid mixture is determined according to the relative mobilities of the two phases, and relative permeability is equal to phase saturation in the borehole.

(5) It is assumed that the two-phase flow behavior follows one of the standard characteristic curves (van Genuchten, Brooks-Corey, or Grant), and that the type of the model is known or can be determined from the available data. Part of the task is to conceive a test design that allows one to distinguish between the different models.

The following model calculations and the sensitivity analysis are based on the supposition that the data to be collected will show, on average, a behavior which can be reproduced by a simplified model based on the above mentioned assumptions. Any deviation from these assumptions will enhance the ambiguity of the solution, increase estimation error, reduce parameter identifiability, and jeopardize the successful analysis of the test as such. Given this fact, the conclusions drawn from this study have to be considered optimistic, i.e. the actual results will be less conclusive because test conditions will not be optimal.

3.3 Error Analysis

The approach used in this study to design an appropriate test is to analyze the sensitivity coefficients, the standard deviations of the estimated parameters, and their correlation structure (see also Appendix A). We try to show that the proposed test sequence is superior to the standard procedure (sequence 1), i.e. it leads to higher sensitivity, thus lower parameter uncertainty and reduced parameter correlation.

The Jacobian matrix represents the sensitivity of the model results with respect to each unknown or uncertain parameter. This information can be used to:

(1) detect the parameters that are likely to be determined with a certain degree of confidence (the test design has to be optimized such that the parameters of interest are associated with large sensitivity coefficients),

(2) identify test events and parts of test events that contain information about the parameters of interest (e.g., increasing sensitivity coefficients with time near the end of a test event suggest to continue measuring; small sensitivity coefficients indicate that the event could be omitted, etc.).

The *a posteriori* covariance matrix of the estimated parameters provides the following information:

(1) It measures the potential accuracy with which a parameter can be determined, provided that the actual system response corresponds to the modeled one, and that the match between the observed and computed pressures is as good as assumed by the prior error variance.

(2) The relative reduction of parameter uncertainty as a result of an improved test design can be evaluated.

(3) It shows the correlations among all the parameters of interest. Note that the uncertainty of one parameter affects the uncertainty of all parameters that are correlated.

(4) It indicates which parameter should be determined by independent measurement.

This kind of analysis is a powerful tool which has many advantages over conventional sensitivity studies. However, the following limitations and restrictions have to be kept in mind (most of which also apply to a standard sensitivity analysis):

(1) The Jacobian matrix is evaluated at a given point in the parameter space (the base case parameter set), i.e. sensitivity coefficients and correlation structure may vary if the actual conditions are different from the assumed ones. This means that the analysis should be repeated for a number of potential parameter combinations.

(2) The error analysis is based on a linearity assumption, i.e. the model output is assumed to vary linearly within the expected parameter variation, which is expressed by its standard deviation (or a multiple thereof). This assumption may be violated if strong two-phase flow effects are present.

(3) Potential parameter variations and data uncertainties have to be provided.

(4) Since the analysis is performed prior to testing, no data are available to actually compute estimation errors and final residuals. The analysis assumes that the average system behavior can be well reproduced by the model, and that the final residuals display the assumed variances.

(5) Most important, the error analysis only deals with *statistical* errors, not with *systematic* ones. The latter may be more important and difficult to account for. Potential errors in the model conceptualization (e.g. wrong type of characteristic curves) have to be analyzed explicitly.

This discussion reveals that the sensitivity analysis provides only semi-quantitative, i.e. relative measures. Nevertheless, they point towards aspects of the test design that can be improved.

In order to address the problem of systematic errors, we will examine the ability of each test sequence to detect errors in the model conceptualization (see Section 5.4 and Appendix A).

4. DEFINITION OF REFERENCE CASE

A base case is defined, representing hydrological unit B of the Valanginian marl at Wellenberg, Switzerland. The parameter set including interval specifications are summarized in Table 2. The permeability of 10^{-16} m^2 corresponds to a hydraulic conductivity of 10^{-9} m/s and a transmissivity of 10^{-8} m^2/s. An initial formation gas pressure of 4200 kPa and a water pressure of 4000 kPa is chosen, approximately representing hydrostatic conditions. We examine different characteristic curves given by the models of van Genuchten (Eq. 1) and Brooks-Corey (Eq. 2), and a third type which is identical to the Brooks-Corey model except that the gas relative permeability is given by $k_{rg} = 1 - k_{rl}$; we will refer to this function as the Grant model. The scaling parameter $1/\alpha$ and p_e, respectively, are chosen such that they give rise to a capillary pressure of 4200 - 4000 = 200 kPa at the initial gas saturation of 0.1. Due to the analogy between these two parameters at low liquid saturations ($1/\alpha \approx p_e$), we will refer to both of them as the *air entry pressure (AEP)*, even though they are numerically different and no finite capillary pressure at $S_l = 1$ is defined in van Genuchten's model. In van Genuchten's model, however, gas is assumed to be immobile for $S_g < S_{gr} = 0.01$, giving rise to an equivalent air entry pressure (defined as the gas overpressure required to displace water) of $p_c(S_g = S_{gr}) = -88$ kPa. Similarly, we relate the van Genuchten parameter n to the *pore size distribution index (PSDI)* λ of Brooks-Corey's model ($n \approx \lambda + 1$). The initial gas saturation of 0.1 is relatively low, as expected at Wellenberg; higher gas contents will be easier to detect. The residual liquid saturation is a parameter that is not likely to be of critical importance as long as the gas saturation remains low. The residual gas saturation, on the other hand, is a sensitive parameter especially for the Grant model since it strongly affects the gas relative permeability at low gas saturations. Finally, the test zone compressibility is assumed to be $2 \cdot 10^{-9}$ Pa^{-1}. Note that this parameter only reflects the compliance of the test system. Liquid and gas compressibility is explicitly taken into account in the model as a function of pressure and temperature. The wellbore storage coefficient as measured in the field at the beginning of a pulse injection test is a combined measure of system and fluid compressibilities.

The van Genuchten model is given by [*Luckner et al.*, 1989]:

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

$$k_{rl} = S_e^{\eta} \cdot \left[1 - (1 - S_e^{1/m})^m\right]^2 \tag{1b}$$

$$k_{rg} = (1 - S_e)^{\gamma} \cdot \left[1 - S_e^{1/m}\right]^{2m} \tag{1c}$$

where

$$S_e = \frac{S_l - S_{lr}}{1 - S_{lr} - S_{gr}} \qquad (S_{lr} < S_l < 1 - S_{gr}) \tag{1d}$$

$$m = 1 - 1/n \tag{1e}$$

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

$$p_c = -p_e \left(\frac{S_l - S_{lr}}{1 - S_{lr}}\right)^{-1/\lambda} \tag{2/3a}$$

$$k_{rl} = S_e^{\frac{2 + 3\lambda}{\lambda}} \tag{2/3b}$$

$$k_{rg} = (1 - S_e)^2 \cdot \left(1 - S_e^{\frac{2 + \lambda}{\lambda}}\right) \tag{2c}$$

For the model referred to as Grant's curves we use the Brooks-Corey functions and apply

$$k_{rg} = 1 - k_{rl} \tag{3c}$$

Table 2: Base case parameter set and interval specifications

Parameter	Value
log (permeability k [m^2])	-16.00
initial formation gas pressure p_i [bar]	42.00
initial formation liquid pressure p_i [bar]	40.00
borehole history pressure p_{bh} [bar]	41.20
air entry pressure: AEP [bar]	
Brooks-Corey p_e	1.72
van Genuchten $1/\alpha$	3.22
pore size distribution index: PSDI [-]	
Brooks-Corey λ	2.00
van Genuchten n	3.00
pore connectivity parameter [-] (van Genuchten):	
η	0.50
γ	0.33
initial gas saturation S_{gi} [-]	0.10
residual liquid saturation S_{lr} [-]	0.25
residual gas saturation S_{gr} [-]	0.01
porosity ϕ [-]	0.02
temperature [°C]	25.00
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
shut-in volume [m^3]	0.10
mud density [kg/m^3]	1050.00

5. RESULTS

5.1 Introduction

In the absence of any measured data and due to the qualitative nature of test design in general (see remarks at the beginning of Section 2), we intend to discuss the results of this analysis in a semi-quantitative manner by looking at the *relative* merit of adding test sequences 2 and 3 to the standard sequence. The measures of comparison are the relative reduction of parameter uncertainty and the change of parameter correlation which will be expressed in terms of a ratio χ between the standard deviations of the conditional and the joint probability density function of the estimated parameters (see Appendix A for details). A low value indicates that the parameter estimate will be ambiguous due to its correlation with other uncertain parameters. If χ approaches 1.0, the corresponding parameter can be determined almost independently.

All these measures are a function of type and quality of the data. We consider the case where only interval pressure data will be available. We further assume that the measurement error is constant throughout the test. This may favor test sequence 1 where in fact stronger pressure fluctuations and thus less reliable data have to be expected. It also assumes that the pressure perturbation at the first occurrence of free gas in the interval actually takes place and will be accurately measured. Note that if this sensitive data is not registered, the potential merit from performing additional test sequences will be more pronounced. On the other hand, we do not consider the gas flow measurements that may be acquired during the RW event. They are expected to fluctuate considerably.

The covariance matrix will be evaluated for all three models (Brooks-Corey, van Genuchten, and Grant) and for all three test sequences. Again, this analysis assumes that the model type is known *a priori*, and that the parameter values are close to the ones specified in Table 2. For this parameter combination, the three models show quite different pressure responses (see Figure 2 below), indicating that choosing the right model is crucial in order to avoid biased estimates. In order to address this potential difficulty, we also examine the ability of the different test sequences to distinguish among alternative models. This procedure will be discussed later in Section 5.4.

5.2 System Behavior

Figure 2 shows the simulated interval pressure for the parameters of Table 2. The three curves are obtained using three different characteristic curves. As a reference, the pressure response under fully liquid-saturated conditions (assuming van Genuchten's curves during the gas injection test) is also plotted. The system behavior can be described as follows:

The slight over-pressure during borehole history displaces the initial mixture of gas and liquid away from the borehole. During the PSR period, pressures start to recover toward the initial liquid pressure. The constant rate pumping test induces strong pressure reductions for the van Genuchten and Brooks-Corey model because the liquid relative permeability is reduced by the presence of free gas. On the other hand, significantly more gas is produced if using the Grant model with its high relative gas permeability, leading to an increased total mobility of the gas-water mixture and thus less pressure reduction during the pumping period. The recovery period is dominated by the gas phase mobility and the total amount of gas trapped in the interval. All curves tend to approach the liquid formation pressure. The pressure increase during the water injection test mostly reflects the amount of free gas in the vicinity of the borehole which in turn is determined by the characteristic curves and the borehole history. Again, the pressure recovers toward the formation liquid pressure. Finally, the gas injection test is most sensitive to the effective gas permeability. Note that the high gas permeability for the Grant model leads to a pressure drop even during the injection period, invoked by the opening of a continuous flow path to the outer gas-bearing zone. If a free gas phase is present, the recovery pressure approaches the gas formation pressure.

Figure 3 shows the gas saturation in the interval. Note that gas enters the borehole already during the PSR period if using the van Genuchten model. The finite air entry pressure of the Brooks-Corey and Grant model lead to a more piston-like displacement of gas during drilling and borehole history. Gas appears thus much later in the borehole. Due to the high relative gas permeability in Grant's model, the gas-water ratio of the produced fluid is likely to be very high even for low gas contents.

We have also looked at the case where no free gas phase is present in the formation, but gas is dissolved at a mass fraction of $5.5 \cdot 10^{-4}$ (mass of methane per mass of liquid), which is slightly below the bubbling pressure under in-situ conditions. It is interesting to note that only limited degassing occurs as liquid is produced. This is due to the fact that water with no dissolved gas is injected during the borehole history period, creating a zone around the well that contains almost de-aired water. A pressure drop within this zone does not invoke degassing. Farther away, however, degassing occurs in a ring-shaped region at a certain distance from the well. The extent of the two-phase region slightly expands with time and moves toward the well. However, only very small gas saturations are achieved, thus not affecting liquid flow. The phenomena are described in detail in *Finsterle* [1994]. For the given parameter combination and provided that only dissolved gas is present, no free gas evolves in the interval under downhole conditions. Gas may still be observed at the surface due to degassing during uplift in the borehole.

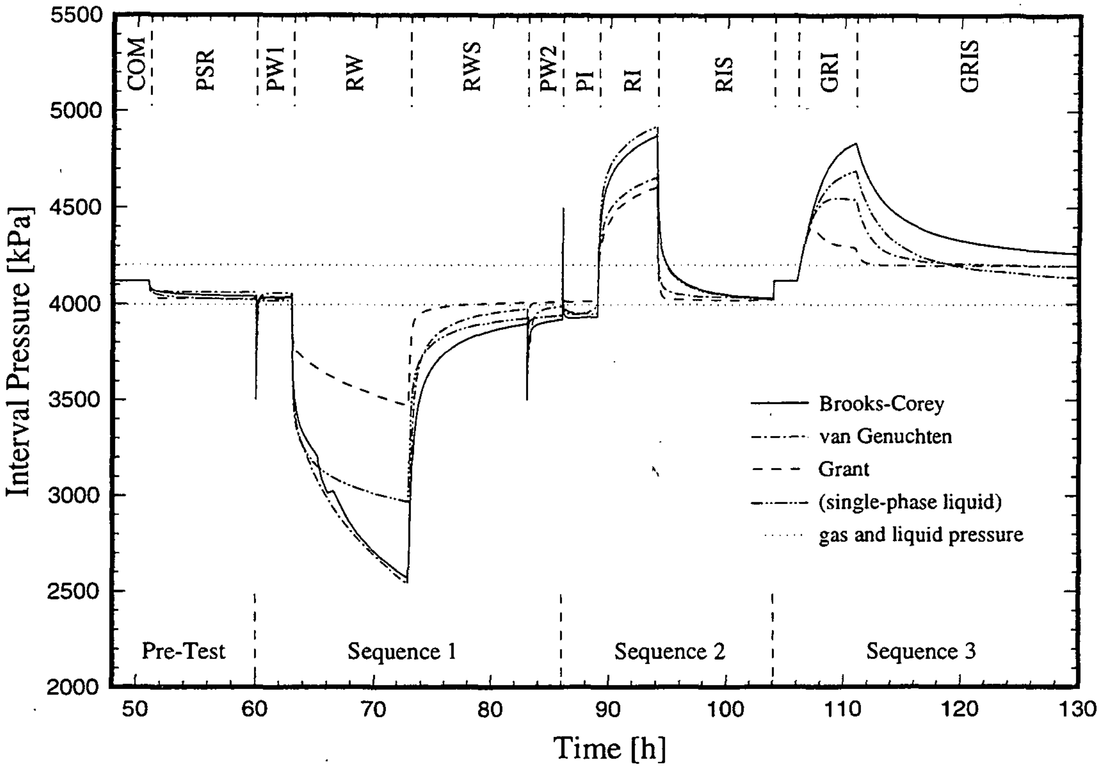

Figure 2: Interval pressure as a function of time

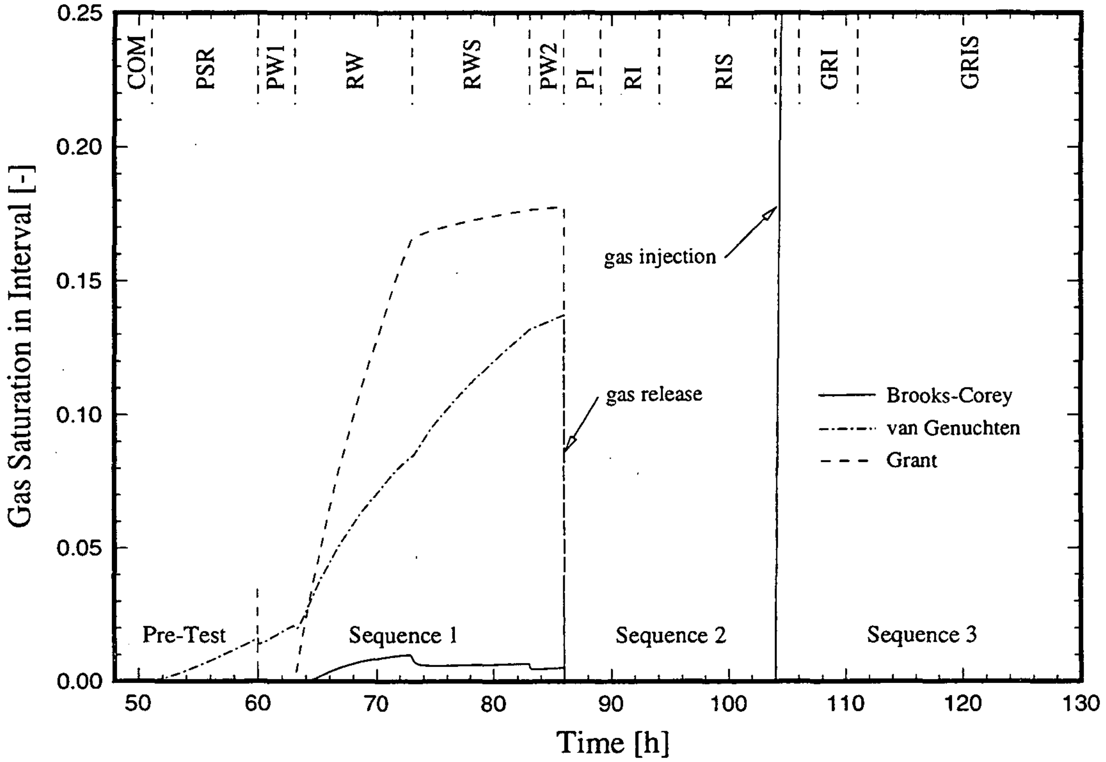

Figure 3: Gas saturation in interval as a function of time

5.3 Benefit From Performing Additional Test Sequences

We have calculated the covariance matrix for each model, first assuming that only data from test sequence 1 are available. Subsequently, data from test sequences 2 and 3 are added to study the reduction of estimation errors and the change of the correlation structure. The result can be summarized as follows:

Tables 3, 4, and 5 demonstrate that performing test sequence 1 alone is insufficient to identify a number of two-phase flow parameters, including the initial gas saturation. This conclusion is based on the high standard deviations calculated for these parameters, i.e. the very large uncertainty prevents one from identifying the parameter value itself. Adding test sequences 2 and 3 considerably improves the determination of the parameters by reducing both the estimation errors and the correlations among the parameters. This conclusion is valid for all three models investigated.

Performing test sequence 2 greatly reduces the standard deviations of all estimates which is basically due to a reduction of parameter correlations. This is reflected in an increase of the ratio χ for most of the parameters.

Figure 4 gives a visual impression of how correlations are reduced as more data become available. In this Figure, each pair of parameters is connected by a line, which also indicates the correlation coefficient on the horizontal axis. The parameter with the highest χ-value is at the top, the one with the lowest χ-value at the bottom. Concentrating on the vertical lines, the pyramids shown in Figure 4 become narrower if more test sequences are added, reflecting a decrease in overall correlation. Note that a physical explanation for parameter correlations is not always easy to find because the impact of indirect correlations has to be taken into account. For example, the positive correlation between gas saturation and permeability has an indirect contribution through their correlation to the formation gas pressure. Gas saturation is usually positively correlated to the formation pressure (the higher the pressure, the more gas is needed to keep storativity constant, which also reflects the pressure dependency of gas compressibility). The formation pressure is slightly positively correlated to permeability (which is itself a result of indirect correlations). While pairs of parameters may exhibit preferred correlations (i.e. either positive or negative), these structures do not always prevail because they are affected by indirect correlations. Furthermore, value and sign of the correlation coefficients are determined by the type of observations. This stresses the importance of obtaining the covariance matrix. In standard sensitivity analysis, the impact of parameter correlations is usually neglected.

Looking at the key parameters of interest, it is primarily the more independent and more accurate estimate of absolute permeability which allows identification of the initial gas content. During the pumping period, the absolute permeability is positively correlated to the gas saturation, i.e. similar interval pressures are achieved by increasing or reducing both permeability and gas saturation. The strong correlation prevents an independent and accurate estimate of both parameters. The water injection test (sequence 2) provides a more precise estimate of absolute permeability, since the region around the borehole is essentially liquid saturated. At the same time, the estimate becomes less dependent on other parameters, especially the gas content, which can now be determined. Similarly, test sequence 3 is designed to obtain an independent estimate of the air entry pressure (notice the increase of χ for the air entry pressure between sequences 2 and 3). 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 air entry pressure is estimated.

Grant's model assumes a high gas mobility even for low gas saturations. It is mainly this peculiarity of the model that explains why gas-related parameters (p_e, λ, S_{lr}, and S_{gr}) are more sensitive and thus much easier to determine compared to the alternative models. If the formation behaves according to the van Genuchten or Brooks-Corey model, it remains questionable whether reliable estimates for this group of parameters can be obtained. As an alternative approach, one might perform a separate test designed to determine two-phase flow parameters that are relevant to performance assessment studies. This test may be a combination of sequences 2 and 3, however with prolonged injection periods. A longer water injection period ensures that the formation is almost completely water saturated prior to starting the gas threshold pressure test. The latter should be extended far beyond the initial gas breakthrough to obtain data in a wider range of saturations and capillary pressures. The design of such a test is not part of the current study, but could be approached using similar techniques.

Table 3: Reduction of estimation error and parameter correlation as a result of adding test sequences 2 and 3, Brooks-Corey model

Brooks-Corey						
	Sequence 1 only		Sequence 1 and 2		Sequence 1, 2 and 3	
Parameter	σ	χ	σ	χ	σ	χ
log (k [m^2])	0.20	0.05	0.04	0.22	0.03	0.30
p_i [bar]	1.25	0.13	1.02	0.12	0.32	0.34
S_{gi} [-]	N/D	0.00	0.07	0.09	0.05	0.13
AEP p_e [bar]	N/D	0.13	N/D	0.11	0.36	0.31
PSDI λ [-]	N/D	0.01	N/D	0.08	1.79	0.23
S_{lr} [-]	N/D	0.01	N/D	0.05	0.45	0.10
S_{gr} [-]	N/D	0.01	0.04	0.19	0.03	0.27
ϕ [-]	N/D	0.01	0.01	0.39	0.01	0.41
log (c_{tz} [Pa^{-1}])	0.43	0.82	0.24	0.95	0.24	0.96
p_{bh} [bar]	0.53	0.27	0.48	0.29	0.25	0.40

σ : standard deviation from joint probability density function
χ : ratio of conditional and joint standard deviation; low value indicates high correlation
N/D : not detectable, i.e. standard deviation much larger than parameter value

Table 4: Reduction of estimation error and parameter correlation as a result of adding test sequences 2 and 3, van Genuchten model

	van Genuchten					
	Sequence 1 only		Sequence 1 and 2		Sequence 1, 2 and 3	
Parameter	σ	χ	σ	χ	σ	χ
log (k [m^2])	0.32	0.04	0.03	0.26	0.03	0.27
p_i [bar]	N/D	0.01	0.64	0.19	0.21	0.50
S_{gi} [-]	N/D	0.03	0.03	0.12	0.02	0.14
AEP $1/\alpha$ [bar]	N/D	0.02	0.52	0.14	0.26	0.27
PSDI n [-]	N/D	0.01	0.87	0.05	0.43	0.09
S_{lr} [-]	N/D	0.02	0.16	0.06	0.10	0.10
S_{gr} [-]	N/D	0.02	0.01	0.06	0.01	0.07
ϕ [-]	N/D	0.01	0.01	0.07	0.01	0.11
log (c_{tz} [Pa^{-1}])	N/D	0.29	0.19	0.09	0.19	0.09
p_{bh} [bar]	0.62	0.19	0.25	0.45	0.22	0.51

σ : standard deviation from joint probability density function
χ : ratio of conditional and joint standard deviation; low value indicates high correlation
N/D : not detectable, i.e. standard deviation much larger than parameter value

Table 5: Reduction of estimation error and parameter correlation as a result of adding test sequences 2 and 3, Grant model

	Grant					
	Sequence 1 only		Sequence 1 and 2		Sequence 1, 2 and 3	
Parameter	σ	χ	σ	χ	σ	χ
log (k [m^2])	0.22	0.10	0.04	0.39	0.04	0.39
p_i [bar]	1.23	0.08	0.11	0.72	0.09	0.81
S_{gi} [-]	N/D	0.08	0.03	0.32	0.03	0.34
AEP p_e [bar]	1.14	0.08	0.02	0.91	0.02	0.91
PSDI λ [-]	0.97	0.61	0.91	0.59	0.61	0.75
S_{lr} [-]	N/D	0.51	0.14	0.65	0.13	0.62
S_{gr} [-]	N/D	0.19	0.01	0.57	0.01	0.62
ϕ [-]	0.01	0.68	0.01	0.52	0.01	0.57
log (c_{tz} [Pa^{-1}])	0.40	0.94	0.27	0.98	0.23	0.96
p_{bh} [bar]	0.41	0.46	0.33	0.54	0.32	0.55

σ : standard deviation from joint probability density function
χ : ratio of conditional and joint standard deviation; low value indicates high correlation
N/D : not detectable, i.e. standard deviation much larger than parameter value

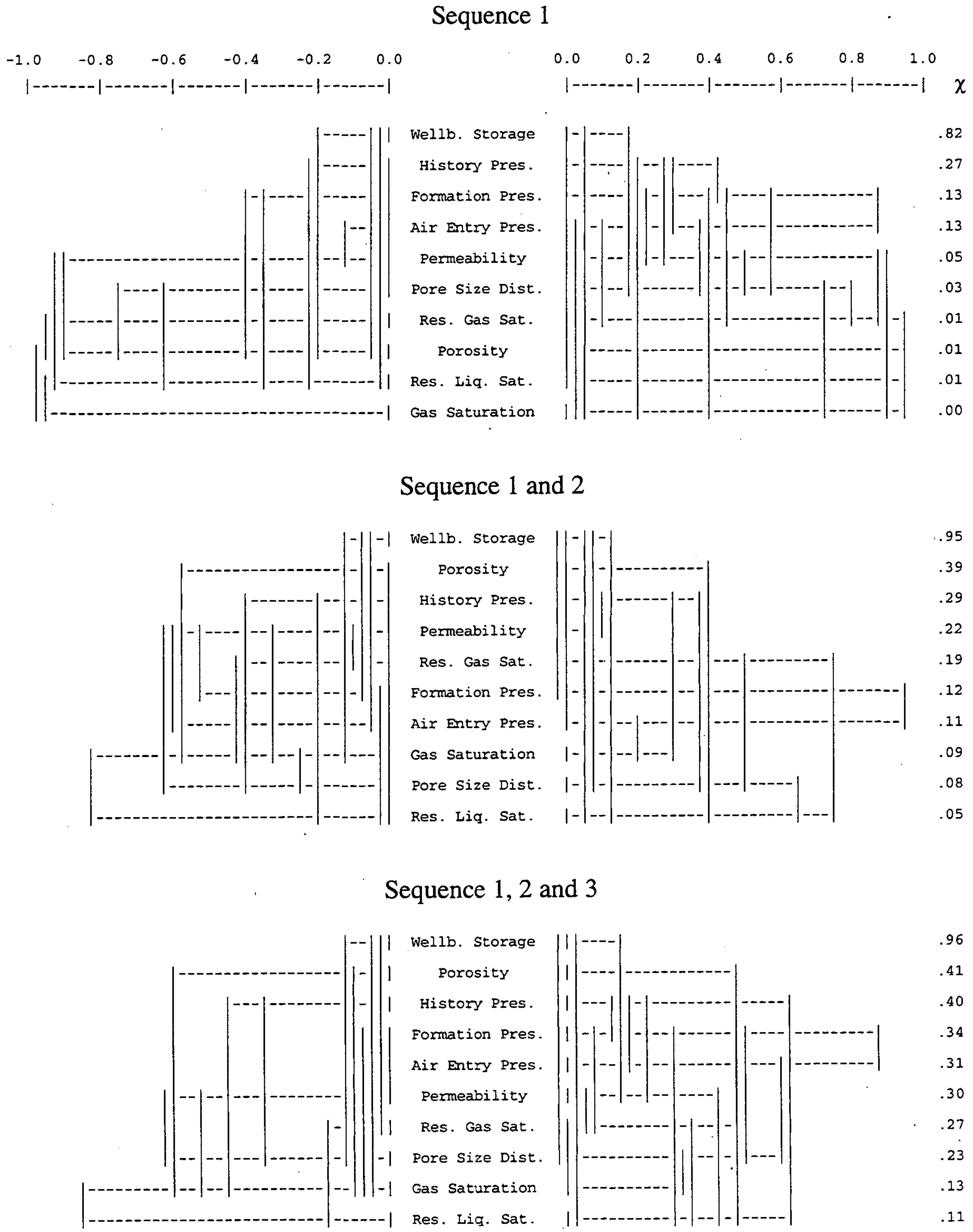

Figure 4: Visualization of correlation structure, Brooks-Corey model

5.4 Model Selection

As pointed out in Section 5.1, it is crucial to examine whether a certain design is able to identify a more likely conceptual model by rejecting its competing alternatives. Even though the three different characteristic curves considered here may lead to very different interval pressures for the given parameter set (see Figure 2), it cannot be excluded that essentially the same response could originate from using either of the models. Previous studies have indicated [*Senger and Jaquet*, 1994; *Finsterle*, 1994] that the origin of the gas observed at the wellhead can only be assessed if the functional form of the characteristic curves is known. Moreover, if the wrong model is chosen, the estimated parameters will be biased despite the fact that a good match is obtained.

The procedure adopted here is to generate synthetic data based on van Genuchten's model with added random noise. Subsequently, each of the three conceptual models is allowed to fit the data using inverse modeling techniques. The estimated error variance will be calculated, representing the goodness-of-fit. As shown in Appendix A, this measure can be used to examine the ability of the test design to discriminate among alternative conceptual models. Since the data have been generated based on van Genuchten's functions, it is clear that using this model for the inversion will show the best performance. The main purpose of the procedure is therefore to reveal how strong the other models are rejected, if at all, and whether adding test sequences 2 and 3 helps identify the model that is most likely to represent formation conditions.

Figures 5 through 7 show the synthetically generated pressures, and the match that was obtained by fitting the three different models to the data. The goodness-of-fit criteria are summarized in Table 6. Note that since the estimated error variance is a random variable, a statistical test can be applied to decide whether one model performs significantly better than its competing alternative. The quantile of the F-distribution on a confidence level of 95% is 1.35, i.e. if the ratio of two error variances is larger than 1.35, then the model with the lower value provided a significantly better fit to the data. For details see Appendix A.

We notice that - besides the van Genuchten model - also the Grant model is able to accurately fit the data if only test sequence 1 is performed. This is consistent with the finding of *Finsterle* [1994] who showed that pressure and flow rate data from a similar test sequence can be equally well matched using grossly different conceptual models. However, if test sequences 1 and 2 are inverted simultaneously, only van Genuchten's model reaches the prescribed standard deviation of 50 kPa, as expected. The two alternatives perform significantly worse, especially the Grant model. Adding test sequence 3 allows one to even more clearly distinguish between the van Genuchten and the Brooks-Corey model due to the fact that the latter exhibits a finite air entry pressure which is laid open during the GRI event.

It should be realized that this analysis does not favor any of the models, i.e. one cannot conclude that the Grant model is less likely to explain field conditions. The statement about model performances is also less conclusive if real data are used which are not derived from one of the mathematical models discussed here. However, this analysis clearly shows that model identification is impossible if only sequence 1 is performed, and that the chances to identify the true model increase substantially if test sequences 2 and 3 are added.

Table 7 shows the estimated parameter values as a result of the inversion. When using the van Genuchten model, the true values are estimated if detectable at all. These values would be accepted by a modeler since the van Genuchten model performs significantly better than the competing alternatives. The one important exception is the Grant model which is able to accurately reproduce test sequence 1. The estimates from this inversion, however, deviate from the true values. Since test sequence 1 is not selective with respect to the conceptual model, these wrong values seem equally acceptable than the ones estimated with the van Genuchten model. This explains the inconclusive results obtained in the previous studies by *Senger and Jaquet* [1994] and *Finsterle* [1994].

Table 6: Estimated error variances using different characteristic curves; data are generated using the van Genuchten model, perturbed by a normally distributed noise with σ_0 = 50 kPa.

Estimated Error Variance s_0^2/σ_0^2			
Model	Sequence 1 only	Sequence 1 and 2	Sequence 1, 2 and 3
data (van Genuchten)	1.00	1.00	1.00
van Genuchten	0.92	0.99	1.06
Brooks-Corey	*3.52*	*2.81*	*4.39*
Grant	0.95	*5.77*	*4.78*

s_0^2 : estimated error variance = *a posteriori* error variance of residuals;
if s_0^2/σ_0^2 is greater than $F_{0.95}$=1.35 (*italic*), the model is unlikely to explain the data

Table 7: Parameter estimates obtained by fitting the wrong conceptual model to data generated using the van Genuchten model. The Grant model seems acceptable if only test sequence 1 is conducted.

	true	Sequence 1 only		Sequence 1 and 2		Sequence 1, 2 and 3	
Param.	value	BC	Grant	BC	Grant	BC	Grant
log (k)	-16.00	*-15.99*	-16.41	*-16.07*	*-16.06*	*-15.89*	*-16.03*
p_i	42.00	*42.27*	42.69	*42.44*	*41.38*	*41.71*	*41.27*
S_{gi}	0.10	N/D	0.05	*0.25*	*0.01*	*0.11*	*0.02*
AEP p_e	1.71	N/D	2.19	*2.16*	*1.86*	*1.38*	*1.68*
PSDI λ	2.00	N/D	N/D	N/D	N/D	N/D	N/D
S_{lr}	0.25	N/D	N/D	N/D	N/D	N/D	N/D
S_{gr}	0.01	N/D	N/D	N/D	*0.01*	N/D	*0.01*
ϕ	0.02	N/D	N/D	N/D	N/D	N/D	N/D
log (c_{tz})	-8.70	N/D	N/D	*-8.70*	*-8.70*	*-8.70*	*-8.70*
p_{bh}	41.20	N/D	N/D	N/D	N/D	N/D	N/D

italic : model is rejected, i.e. the van Genuchten model is recognized as the better alternative
N/D : not detectable, i.e. standard deviation much larger than parameter value

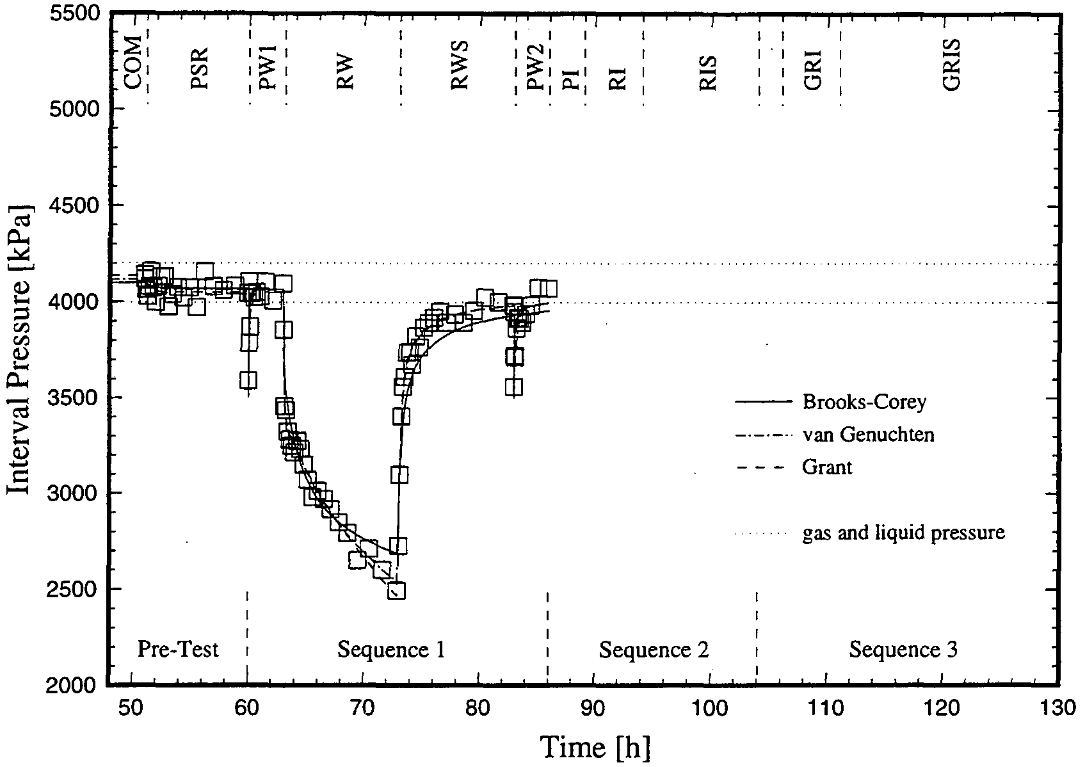

Figure 5: Fitting alternative models to data synthetically generated for test sequence 1

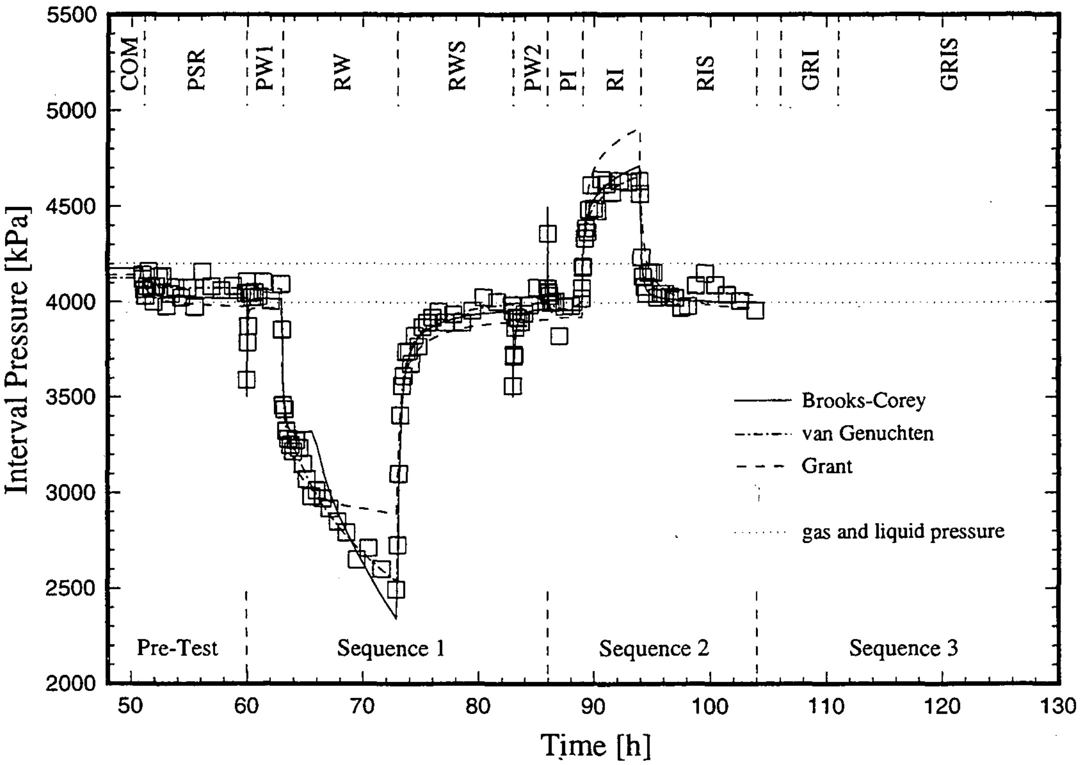

Figure 6: Fitting alternative models to data synthetically generated for test sequences 1 and 2

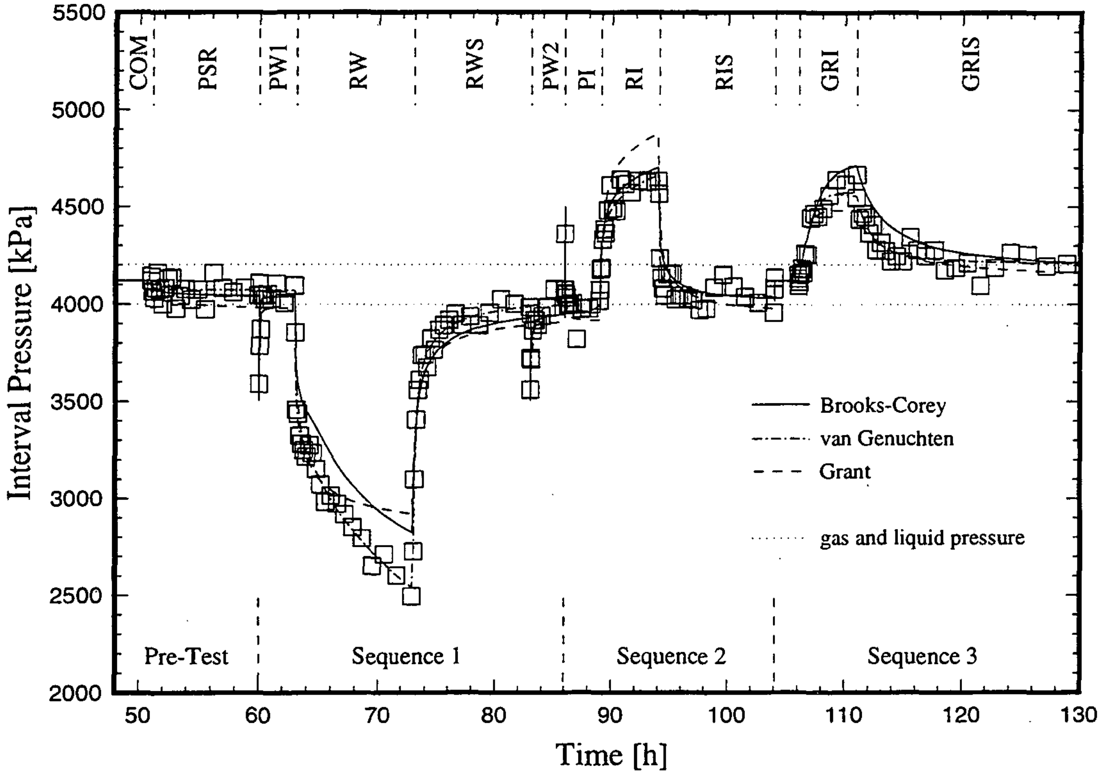

Figure 7: Fitting alternative models to data synthetically generated for test sequences 1, 2 and 3

6. CONCLUDING REMARKS

6.1 Summary

A test sequence has been designed for a low permeability, low porosity formation with potentially small contents of natural gas. The base case parameter set represents hydrological unit B of the Valanginian marl at the Wellenberg site in Switzerland. The objectives of the test is to:

(1) assess the presence of a free gas phase,
(2) obtain unbiased permeability and pressure head estimates,
(3) determine two-phase flow parameters for performance assessment studies.

It was recognized from previous studies that performing a standard test sequence consisting of a series of pulse withdrawal and pumping tests leads to ambiguous results mainly due to the fact that

(1) the pressure and flow rate data are noisy because of the emergence of gas bubbles in the interval,
(2) the data is not selective with respect to the conceptual model, i.e. it is not possible to identify the most likely relative permeability and capillary pressure functions,
(3) the parameter estimates are relatively uncertain and highly correlated.

Theoretically, these drawbacks can be overcome by adding a second test sequence consisting of a water injection test. Joint inversion of both test sequences

(1) reduces uncertainties regarding the physical processes occurring in the test zone,
(2) strongly reduces parameter correlations, i.e. allows for more independent parameter estimates,
(3) produces unbiased permeability and improved head estimates,
(4) allows identification of a free gas content,
(5) reduces estimation errors of all parameters.

A third test sequences is proposed consisting of a gas injection test. Joint inversion of all three test sequences

(1) allows identification of gas-related formation parameters, and
(2) further reduces parameter uncertainties.

The analysis described in Section 5.4 shows that

(1) identification of appropriate characteristic curves is impossible if only sequence 1 is performed,
(2) the chances to identify the true model increase substantially if test sequences 2 and 3 are added,
(3) choosing the right conceptual model is crucial to meet the objectives of the test if gas is actually present in the formation.

All these conclusions are of a qualitative nature. They can only be reached if

(1) the field conditions are similar to the base case parameter set summarized in Table 2,
(2) the system can be approximated by a simplified conceptual model,
(3) control over each test event can be guaranteed,
(4) accurate data can be collected representing downhole conditions.

We believe, however, that the analysis conclusively shows the potential benefit of performing additional test sequences.

6.2 Recommendations

6.2.1 Proposed Field Activities

Practical recommendations for field testing are given here. The rationale behind each of the suggestions is not recapitulated.

(1) Pretest activities should be carefully reported. Estimate annulus pressures, temperatures, densities, and fluid losses during the borehole history period. Keep borehole history period as short as possible. Avoid strong perturbations.

(2) Provide detailed reports on test zone configuration (especially interval volume, test zone compressibility, position of pressure and flow rate metering equipment, provide accuracy estimates of equipment). Try to measure or estimate downhole conditions, especially flow rates.

(3) Allow for a long PSR phase.

(4) Perform pulse withdrawal test. Watch for increased test zone compressibility.

(5) Perform a constant rate withdrawal test. The chosen flow rate has to fulfill two essential requirements:

  - flow rate has to be controllable, i.e. measurable with a high degree of accuracy,
  - the flow rate should be as small as possible.

  The first requirement prevails. The duration of the pumping period should be as long as feasible, interval pressures approaching steady state. Watch for gas appearances, and measure gas-water ratio, if possible.

(6) Allow for a long PWS period. Watch for pressure fluctuations.

(7) Perform second pulse withdrawal test. Determine whether the interval contains free gas.

(8) Perform quick look analysis, i.e. draw diagnostic plots for all events. Watch for any anomaly in the system behavior. Analyze data for each test event separately using standard single-phase evaluation techniques (type curves, wellbore simulator). Check whether late time data can be exactly matched, and whether consistent estimates for permeability, storativity, and pressure head are obtained.

(9) Decide whether the second test sequence shall be invoked. Performing the second test sequence is strongly recommended if there is any evidence for free gas in the formation, such as:

- gas is produced at the surface at a fairly high gas-water ratio,
- free gas is accumulated in the interval,
- data cannot be matched using standard single-phase interpretation techniques,
- the diagnostic plots indicate composite system behavior,
- pressure and flow rate fluctuations occurred during pumping.

(10) Release all the gas that is accumulated in the test interval. This may require deflating the upper packer and flushing all the lines. A short compliance period should be added.

(11) Perform pulse injection test. Determine system compressibility. No free gas should be trapped in the interval.

(12) Perform constant rate water injection test RI. The flow rate should be as small as possible, nevertheless accurately measurable. Continuously draw diagnostic plots. Terminate injection as soon as wellbore storage effects ceased and a reliable estimate of permeability could be obtained.

(13) Allow for a relatively long RIS period. Continuously draw diagnostic plots, and watch for composite system behavior.

(14) Decide whether the third test sequence shall be invoked. Performing the third test sequence is strongly recommended if:

- two-phase flow parameters are to be determined,
- accurate estimate of gas content is desired, and
- pressure head estimates from the previous test events are inconsistent.

(15) Displace fluid in interval by gas using control lines.

(16) Perform constant rate gas injection test GRI. Try to identify threshold pressure. Extend injection significantly beyond the threshold pressure.

(17) Allow for a long GRIS period.

(18) Deflate packers.

6.2.2 Alternative Testing

If performing all three test sequences in a row is not considered feasible, the objectives of sequence 3 could be achieved in a separate test. After completion of drilling and initial testing, i.e. if time constraints are less severe, one might go back to a selected location of the borehole. Packers are set as closely as possible to achieve a small interval volume. An largely extended PSR period may be followed by a test similar to sequence 2 and 3 described above. The duration of each event, however, should be much longer. Design calculations have to show that an extended gas threshold pressure test without prior water injection is sufficient to reliably determine two-phase flow parameters. This is certainly true in a single-phase environment.

Another possibility might be to collect additional data in a second observation interval somewhat deeper in the borehole. This provides information about a better defined volume of rock, and may improve the sensitivity of two-phase flow parameters if breakthrough of either gas or liquid can be observed. Again, the situation should be investigated in a separate design study design.

6.2.3 Future Design Calculations

The situation at Wellenberg (low permeabilities, small porosities, free gas at low saturations) requires special efforts in designing, conducting, and analyzing hydraulic tests. First, the physical processes and potential system behaviors have to be understood. This is best done by performing a series of numerical simulations under different initial and boundary conditions. Several studies of this kind have been conducted in the past [*Senger*, 1994; *Senger and Jaquet*, 1994; *Jaquet*, 1994; *Finsterle*, 1994]. While a standard sensitivity analysis provides information regarding the impact of parameter variations on the system response, it does not address the question of whether these parameters can be actually determined based on selected observations of the system state. This is the classical difference between identifiability and non-uniqueness described by *Carrera and Neuman* [1986]. Sections 5.3 and 5.4 of this report clearly demonstrate that design calculations have to be conducted using an inverse formulation. We recommend that this is the methodology to be pursued in future optimization studies.

ACKNOWLEDGMENT

This work was supported, through U.S. Department of Energy Contract No. DE-AC03-76SF00098, by the Director, Office of Civilian Radioactive Waste Management, Office of External Relations, administered by the Nevada Operations Office in cooperation with the Swiss National Cooperative for the Disposal of Radioactive Waste (Nagra). Thanks are due to P. Marschall (Nagra), C. Oldenburg and B. Freifeld (LBL) for a careful review of the manuscript.

GLOSSARY

BH	Borehole history period
COM	Compliance period before shut-in
DEF	Packer deflation
GRI	Constant rate gas injection test
GRIS	Pressure recovery after constant rate gas injection test (shut-in)
INF	Packer inflation
PI	Pulse injection test
PSR	Pressure static recovery (shut-in)
PW	Pulse withdrawal test
RI	Constant rate injection test
RIS	Pressure recovery after constant rate injection test (shut-in)
RW	Constant rate withdrawal test
RWS	Pressure recovery after constant rate withdrawal test (shut-in)

REFERENCES

Afams, 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, February 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.

Enachescu, C., B. Harborth, M.B. Kloska, L. Ostrowski and W. Tölke, WLB: Hydraulic Testing SB3, Interval Report SB3 / VM28 / 1192.99m, Nagra Internal Report, Nagra, Wettingen, Switzerland, June 1992.

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

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

Finsterle, S. , and K. Pruess, Solving the Estimation-Identification Problem in Two-Phase Flow Modeling, *Water Resour. Res.*, 31(4), 913-924, 1995.

Jaquet, O, WLB: Modelling of hydrotest SB4-VM2, sensitivity analysis and estimation of two-phase flow parameters, Nagra Internal Report, Nagra, Wettingen, Switzerland, March 1994.

Lavanchy, J.M. and R.T. Johns: WLB: Gas Production from Hydrotests at Wellenberg, Nagra Internal Report, Nagra, Wettingen, Switzerland, March 1994.

Luckner, L., M.Th. van Genuchten and D. Nierlsen, 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.

Pruess, K.: TOUGH User's Guide, Nuclear Regulatory Commission Report NUREG/CR-4645, NRC, Washington, DC, 1987, [also published as Report No. LBL-20700, Lawrence Berkeley Laboratory, Berkeley, Calif., August 1987].

Pruess, K., TOUGH2 - A General-Purpose Numerical Simulator for Multiphase Fluid and Heat Flow, Report No. LBL-29400, Lawrence Berkeley Laboratory, Berkeley, Calif., May 1991.

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

Senger, R.K., Preliminary Results of Design Calculations for Hydro Testing Under Two-Phase Conditions UGIV Task 120, unpublished report, Colenco Power Consulting Ltd., Baden, Switzerland, November 1994.

APPENDIX A: A NEW APPROACH TO DESIGN CALCULATIONS

In this Appendix we discuss some of the statistical measures which are used to examine the performance of a test design. We presume that the objective of such a test is to determine parameter values which characterize the hydraulic properties (e.g. absolute permeability) or the in-situ conditions (e.g. gas saturation) of the formation. Thus, the design elements of a test (i.e. type, duration, and sequence of test events, type of data to be measured, location of observation points, etc.) should be chosen such that the uncertainty of the estimated parameters is acceptably small. Optimization of a test design can be described as an iterative process consisting of the following steps:

(1) Conceive a test design, i.e. sequence of test events, type of data to be collected, location of sensors, number and type of parameters to be identified, etc.

(2) Generate synthetic data for all potential observation points. Measurement errors can be reproduced either by assigning a certain standard deviation to the data points, or by explicitly perturbing the data by adding random noise.

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

(4) Analyze the Jacobian matrix which provides information regarding the sensitivity of each data point with respect to each parameter. Revise the test design to improve the sensitivity of the data.

(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 e.g. use different characteristic curves to examine the impact of the conceptualization on the parameter estimates. Try to fit alternative models to the data generated in step 2. Check the overall fit. If the data 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 may be biased.

In this Appendix we provide the theoretical background of the statistical measures being used to examine the performance of a test. In order to simplify the discussion, we assume that only one type of data will be measured. The calculated system response (e.g. pressure measurement at a certain point in space and time) will be referred to as z_i; parameters are designated with p_i.

First of all, estimates of measurement errors as well as potential variations of the model parameters have to be specified. Since no actual measurements are available yet, these estimates have to be based on experience, thus introducing a somewhat subjective element into the design calculations. One should realize, however, that this kind of judgment is common practice for any standard sensitivity analysis, where parameters are altered by a certain amount which is believed to represent the potential parameter variation, and where the resulting deviation of the calculated system response is compared to the expected measurement error in order to decide whether it is significant or not. Furthermore, as will be shown here, the impact of these subjective estimates can be largely eliminated by normalization. Finally, in many cases it is sufficient to compare the *relative* performance of competing test designs. This further reduces the impact of subjective decisions.

The potential measurement error, assumed to be Gaussian, can be described by a covariance matrix **C**, with the variances of the measurements on its diagonal. Note that only the *relative* size of the variances 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 **V**, where $\mathbf{V}^{-1}$ is the weighting matrix to be used for the solution of the inverse problem:

$$\mathbf{C} = \sigma_0^2\, \mathbf{V} \tag{A1}$$

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 is given by:

$$s_0^2 = \frac{\mathbf{r}^T \mathbf{V}^{-1} \mathbf{r}}{m - n} \tag{A2}$$

where **r** is the residual vector, containing the differences between observed and calculated pressures, m is the number of measurements, 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 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, then either the model is inappropriate to reproduce the data, or the magnitude of the measurement errors reflected by matrix **C** were underestimated. Note that if design calculations are performed, no data are available and thus no residuals can be calculated. However, expectations regarding the residuals are expressed through matrix **C**, meaning that s_0^2 can simply be replaced by σ_0^2 for the subsequent error analysis.

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

$$\mathbf{C}_p = s_0^2 \left(\mathbf{J}^T\, \mathbf{V}^{-1}\, \mathbf{J}\right)^{-1} \tag{A3}$$

where **J** is the Jacobian matrix of dimension $m \times n$ with elements $J_{ik} = \partial z_i / \partial p_k$. If actual data are analyzed, the Jacobian is 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 field conditions. It is obvious that the results of the design calculations are only valid for this parameter set, i.e. they may change considerably if the actual field conditions are different from the expected ones. Therefore, one might want to repeat the design calculations for a variety of parameter combinations to assess the conclusions. This difficulty also applies to standard sensitivity analysis, of course.

The interpretation of the covariance matrix $\mathbf{C}_p$ provides the key criteria based on which the experimental design can be improved. First we note that $\mathbf{C}_p$ 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 $\mathbf{C}_p$ contain the variances σ_{pi}^2 of the estimated parameters p_i. The test design should be optimized primarily with respect to this measure. The relative comparison between competing designs is straightforward: The test sequence which yields smaller σ_{pi}^2 performs better than its competing alternative since it allows for a more reliable determination of the parameters of interest. However, the question remains whether the absolute value of σ_{pi}^2 is *sufficiently* small. The uncertainty of an estimated parameter can be considered acceptable if it does not lead to inadmissible errors in the subsequent predictive simulations. In order not to overestimate the accuracy of the parameter estimates, it is recommended to choose a conservative value for σ_0^2, i.e. the matrix $\mathbf{C}_p$ should not only reflect measurement errors but also the uncertainty of the conceptual model as well as the uncertainty of all the parameters which are not subjected to the analysis but might be correlated to the parameters of interest.

Next we shortly discuss the impact of correlations on the estimation error. Correlation among parameters can be described as a conjoint impact of parameter changes on the system behavior. For example, if two parameters are negatively correlated, a similar system response or - more precisely - a similar s_0^2-value 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 entities of parameter combinations. They obviously depend on the data available, and also on the number of simultaneously estimated parameters, since indirect correlations may overwhelm the direct correlations (for details see *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 $\mathbf{C}_p$, which are the variances from the joint probability density function, account for this fact. They have to be distinguished from the conditional standard deviation $\sigma_{p_i}{}^*$ which measures the uncertainty of a parameter assuming that all the other parameters are either exactly known or uncorrelated. The conditional standard deviation is obviously always smaller than the one from the joint probability density function. The situation is illustrated in Figure A1 for the case of two parameters. The ratio

$$\chi_i = \frac{\sigma_{p_i}{}^*}{\sigma_{p_i}} \qquad \text{(A4)}$$

is a measure of how independently parameter i can be estimated. Small values of χ_i usually indicate that the uncertainty σ_{p_i} 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 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.

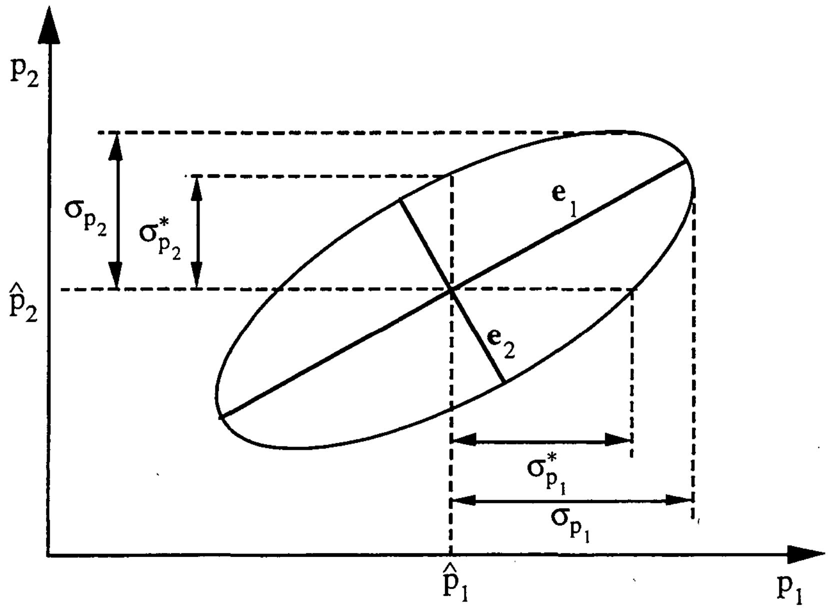

Figure A1: Confidence region, joint and conditional standard deviation

So far we have discussed the criteria based on which the performance of a test design can be determined. No guidelines have been given regarding the optimization process itself, i.e. how the test design could be improved to actually reduce σ_{pi}. No general recommendation can be given. However, the Jacobian matrix **J** (a by-product of the previous analysis) 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:

$$\kappa_{ij} = \frac{\partial z_i}{\partial p_j} \cdot \frac{\sigma_{pj}}{\sigma_{zi}} \tag{A5}$$

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 as follows:

$$\zeta_i = \sum_{j=1}^{n} |\kappa_{ij}| \tag{A6}$$

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 or activated during a different time segment of a transient test, 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.

One might also compare the overall parameter sensitivity π_j

$$\pi_j = \sum_{i=1}^{m} |\kappa_{ij}| \tag{A7}$$

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.

Again, the measures ζ_i and π_j are somewhat subjective because they require specifying values for σ_{pj} and σ_{zi}. However, they point towards aspects of the test configuration that can be modified to improve the overall design.

In the remainder of this Appendix we discuss the use of *model identification criteria.* They are employed here to study the ability of a test to distinguish between different *conceptual* models. For example, data from a hydraulic test under two-phase flow conditions may be used to identify the functional form of the characteristic curves, e.g. whether van Genuchten's, Brooks-Corey's or Grant's model is more appropriate. If each of these models were able to equally well match the data, then the data are *not* selective with respect to the conceptual model. In this case, the estimates are likely to be biased if the wrong model is chosen. Test design should therefore also address the question of model identifiability.

As a general rule, an experiment should be configured such that the salient features of a certain conceptual model are revealed. In other words, a competing model should fail if it lacks some of the characteristics which are pertinent to the system.

The procedure proposed here to test model identifiability consists of several steps. First, synthetic data are generated using a certain conceptual model, and noise is added to simulate measurement errors. Second, a number of alternative conceptual models are developed and tested against the data, i.e. the inverse problem is solved for each of these models. Third, the performance of the inversion is measured by means of so-called *model identification criteria.* Finally, the model identification criteria are statistically tested against each other to see whether the "true" model performs significantly better than the competing alternatives. Since the "true" model is not known, synthetic data have to be generated for all competing alternatives, and the procedure has to be repeated for all test designs under consideration.

While a number of rather sophisticated model identification criteria have been presented in the literature (CARRERA 1984), the qualitative nature of these design calculations justifies the use of a simple measure of goodness-of-fit, namely the estimated error variance s_0^2 (see Equ. A2). The smallest value for s_0^2 is obtained for the "true" model (the one that was used to generate the data). Note that if the variances defined in matrix **C** are consistent with the ones used to perturb the data, s_0^2 will be 1.0. Alternative models will realize larger s_0^2-values due to an imperfection in the fit. Since s_0^2 is a random variable, the ratio $s_0{}^2{}_i/s_0{}^2{}_j$ can be statistically tested. If

$$\frac{s_0{}^2{}_i}{s_0{}^2{}_j} < F_{m-n_i,m-n_j,1-\alpha} \qquad \text{for } s_0{}^2{}_i > s_0{}^2{}_j \qquad \text{(A8)}$$

then model *i* and model *j* perform equally well because the differences in goodness-of-fit are not statistically significant. In (A8), $F_{m-n_i,m-n_j,1-\alpha}$ is the quantile of the F-distribution with the two degrees of freedom of model *i* and *j*, respectively. $1-\alpha$ is the confidence level, with α being the risk that model *i* is considered equal to model *j* even though it is not. A selective test design is one for which the model test always fails if model *j* is the "true" model and model *i* is one of its wrong competitors. Note that since models are only tested against each other or against synthetically generated data with known error structures, subjective decisions regarding matrix **C** affect the outcome of test (A8) only marginally.

In this Appendix, a procedure has been proposed to improve the design of an experiment. The basic idea is to solve the inverse problem using synthetically generated data, and to examine 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 analyzed.

While the proposed procedure exhibits some of the shortcomings inherent in any design calculations, it has tremendous advantages over a standard sensitivity analysis. Taking an inverse perspective allows one to address the questions 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. Finally, these kinds of design calculations also serve as a preparation for the subsequent data analysis.

The procedure outlined here is supported by the ITOUGH2 code [*Finsterle*, 1993].

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL AND ELECTRONIC
INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720