Optimal Input Design—A Method for Enhancing Data Quality

Bjarne A. Foss, SPE, U. of Trondheim

Summary. This paper presents a generic design method for specifying the flow sequence from one or multiple wells. The method seeks to maximize the information content in collected data. It is shown how the information content, defined in precise terms, is a function of the flow sequence. Three examples demonstrate how the proposed method improves data quality, and guidelines for practical use are presented.

Introduction

This paper investigates the use of the optimal input design method for reservoir evaluation. Optimal input design generally is used to construct system input signals that optimize the information content in obtained data with respect to certain prespecified, unknown properties. Such a method can be used to determine the flow-rate sequence of a well test by specifying the time intervals of the main drawdown and buildup periods.

Current analysis techniques vary from manual to sophisticated computerized methods; however, all methods combine prior knowledge about the reservoir (e.g., through a model) with information from actual reservoir data. Increasing the amount of relevant information in the data will, in turn, increase our knowledge about a given reservoir.

Optimal input design methods have received little attention in the petroleum literature. Dogru and Seinfeld,[1] however, do discuss these methods in conjunction with well testing.

Optimal Input Design Method

The theory of optimal input design investigates methods of constructing system input signals that minimize the uncertainty of prespecified, unknown parameters. This can be illustrated by assuming the following functional relationship between a measurable input, q, and a measurable output, y.

$$y=f(\Theta,q) \quad \text{(1a)}$$

$$\text{and} \quad \dim(q)=\dim(y)=\dim(\Theta)=1, \quad \text{(1b)}$$

where Θ is an unknown parameter that uniquely defines the function f. The identifiability of Θ (i.e., the confidence with which Θ can be computed) improves if the sensitivity of the output y with respect to the parameter Θ increases. It can be shown [see Eq. A-5 for the case when $\dim(\Theta)>1$] that the uncertainty in the calculation of Θ, for a given set of measured data $\{(q^1,y^1)\ldots(q^N,y^N)\}$, is

$$\operatorname{var}(\bar{\Theta})=\left[\frac{1}{N}\sum_{n=1}^{N}(\partial y^n/\partial\Theta)^2\right]^{-1}\operatorname{var}(e), \quad \text{(2)}$$

where $\bar{\Theta}=\Theta-\Theta_0$ and Θ_0=the correct parameter value if the following assumptions are valid.

1. The noise on the observed output data is given by

$$y^n_{OB}=y^n+e^n, \quad \text{(3)}$$

where e^n=random (white) noise source with zero mean value and a constant variance and y=exact value.

2. The function f is linear.

3. The function f describes the exact relationship between q and y; hence, modeling errors do not exist.

These assumptions are restrictive and imply that care must be exercised when the above approach is used.

The essence of the optimal input design method is that *the uncertainty in the estimates of Θ depends on the choice of the input* $\{q^1\ldots q^N\}$ because the sensitivities $\partial y/\partial\Theta$ (see Eq. 2) are generally functions of the input q.

Compared with the above example, the problem increases in complexity when the number of unknown parameters increases because the first term on the right side of Eq. 2 becomes a matrix. This case is treated in Appendix A, and the equivalent of Eq. 2 is

$$C_{ma}(\bar{\Theta})=\mathbf{H}^{-1}\operatorname{var}(e) \quad \text{(4)}$$

$$\text{and } \mathbf{H}=\frac{1}{N}\left[\sum_{n=1}^{N}\left(\frac{\partial y^n}{\partial\Theta}\right)\left(\frac{\partial y^n}{\partial\Theta}\right)^T\right], \quad \text{(5)}$$

where $\dim(\Theta)=M$ and $\dim(\mathbf{H})=M\times M$.

Eq. 4 shows that the Hessian matrix, $\mathbf{H}$, is important for determining the uncertainty in the estimates of the unknown parameters. Consequently, this matrix is closely related to the information content with respect to some prespecified, unknown parameter, Θ, in experimental data.

The general optimal input design method may be formulated in a compact form (see Eq. A-8):

$$\min_{q\in Q} g(\mathbf{H}^{-1}), \quad \text{(6)}$$

where Q=admissible inputs and g=the functional associating a scalar with the matrix $\mathbf{H}$ (g: $R^M\rightarrow R$). Hence, we relate a matrix, $\mathbf{H}^{-1}$, to a functional, $g(\mathbf{H}^{-1})$. A common choice is to use the maximum eigenvalue of $\mathbf{H}^{-1}$. See Appendix A for details about choosing g.

Apart from its use in input design, the analysis of the Hessian matrix, with the noise in the observed data (e in Eq. 3), is important in the selection of which parameters to estimate because this analysis gives a quantitative measure of the uncertainty associated with the parameter estimates. Generally, it is better to fix the value of an unknown parameter rather than estimating it with large error bounds, and fixing such a parameter often reduces the uncertainty in the estimates of the other unknown parameters.

Examples 1 and 2: Optimizing the Flow Sequence in Well Tests

Example 1 is a study of the common well-test situation in which only the well pressure and the surface flow rate are monitored. **Fig. 1** shows the single-phase model of the system. The reservoir is assumed to consist of three horizontal layers; however, as shown in **Fig. 2**, only the second layer is perforated. **Table 1** gives the characteristic properties of the reservoir.

The model describing the reservoir and the well pressure is derived by standard discretization with cylindrical geometry. The discretized model is given by

$$\mathbf{A}(\mathbf{k})p^{n+1}=p^n+\vec{b}(k,k_s)\Delta p^{n+1}, \quad \text{(7)}$$

$$\Delta p^n=\beta(k_s)q^n_s, \quad \text{(8)}$$

$$q^{n+1}_s=\alpha(S)q^n_s+[1-\alpha(S)]q^n, \quad \text{(9)}$$

$$\text{and } p^n_w=p^n_{12}-\Delta p^n. \quad \text{(10)}$$

This is a standard single-phase simulator, where $A(k)$=a tridiagonal matrix, $\vec{b}(k,k_s)$=a vector, and p_{12}=the pressure in the gridblock adjacent to the perforation. β, which depends on skin permeability, couples the sandface flow rate, q_s, to the pressure drop past the damaged zone. α is determined by the wellbore storage con-

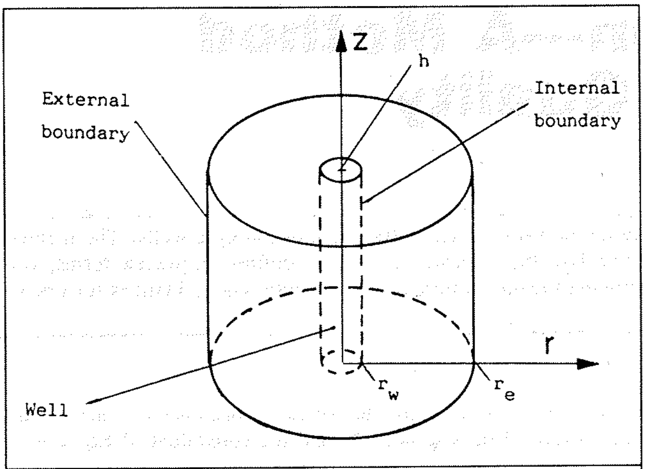


Fig. 1—Cylindrical model geometry.

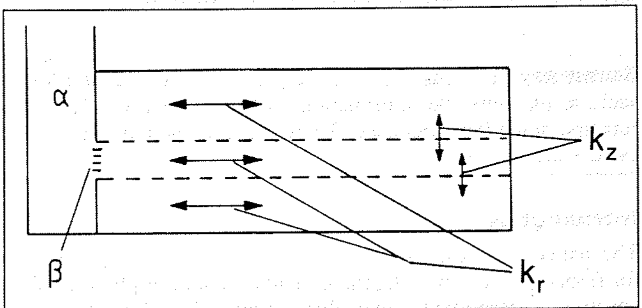


Fig. 2—Vertical cross section.

The optimal input design method is used to construct the flow-rate sequence during a well test. In Example 1, we assume that the testing period consists of one drawdown period and one build-up period and that the time of the total testing period and the flow rate during the drawdown period are fixed. Now, we must choose the instant at which production terminates, as shown in **Fig. 3**. The goal is to select the instant of time that maximizes the information content in the data.

The time variables are chosen to give a total testing time of approximately 2 hours and 45 minutes, with a 25-second sampling time for each of 400 samples. The unknown parameter vector is defined as $\Theta^T=[k_r k_z\ \alpha(S)\ \beta(k_s)]$. Parameter values are assumed to be (see Fig. 2) $k_r=100$ md, $k_z=1$ md, $\alpha=0.10$, $S=2.11\times10^{-7}$ m^3/Pa, $\beta=2.14\times10^8$ $m^3/s\cdot Pa$, and $k_s=116$ md.

In this example, the well is slightly stimulated, but because the method outlined is equally applicable to wells with positive and negative skin factors, the stimulation is inconsequential.

The design criterion, g (see Eq. 6, or for a detailed expression see Eq. A-9), is computed for different switch-off time instants by computing the Hessian matrix with the expression for $\mathbf{H}_2$ in Eq. A-2. The sensitivities, $\partial y^n/\partial\Theta$, are calculated with the equations for the simulator (Eqs. B-1 through B-4). p_w is the output data; hence, it is substituted for y^n in Eqs. A-2 and B-4. **Fig. 4** shows the input design criterion. The curve is constructed by linear interpolation of the computed g values. Note that the F_W value in this and the following examples is based on the guidelines given by Foss.[2]

With only one free variable (t_{sw}) in Example 1, the optimal flow sequence for the well test is determined by visual inspection of Fig. 4. This figure shows how the information content is maximized (i.e., when the performance criterion is minimized) by specification of a test where the flow terminates within 30 to 60% of the total testing time. This implies that the unknown well parameters can be estimated with the least uncertainty if the pressure data from such a well test are used.

Besides the flow sequence, the curve in Fig. 4 depends on the model, the unknown parameters, and the assumed values of these parameters. Choosing values for the unknown parameters presents a problem. Some indication as to the range of values generally is available from seismic and well-log data. This also applies to the choice of the model. An obvious way to circumvent the problem of unknown parameter values is to run several input design computations (similar to that in Fig. 4) for different models and parameter values within the probable range of choices.

A somewhat different situation exists in Example 2 because we assume that an extra set of pressure data is available. **Fig. 5** is a schematic of the production and monitoring equipment and the reservoir layering. The well pressure is monitored by Gauge 1. The pressure in Layer 1 is monitored by Gauge 2. Note that the production string is sealed only at the lower end, permitting a wire to run through to Gauge 2.

Restrictions on the choice of the test flow rate are defined as in Fig. 3, while the time variables are a total testing time of approximately 28 hours, with a sampling time of 4 minutes and 10 seconds for each of 400 samples.

The unknown parameter vector is defined by $\Theta^T=[k_{r2}k_{r3}k_{z1}\ k_{z2}k_{z3}\ \beta(k_s)]$, where $k_{r1}=2k_{r2}$, $k_{r4}=k_{r3}$, and subscripts 1 through 4 identify the permeabilities of the different layers shown in Fig. 5.

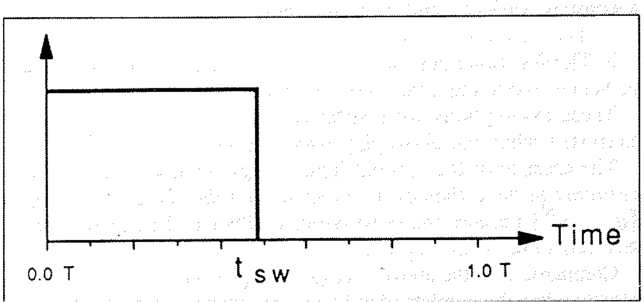
In this example the wellbore storage effect is neglected. Parameter values are $k_{r2}=10$ md, $k_{r3}=50$ md, $k_{z2}=2$ md, $k_{z3}=4$ md, $k_{z4}=20$ md, $\beta=4.28\times10^8$ $m^3/s\cdot Pa$, and $k_s=230$ md.

The curve in **Fig. 6** is computed in the same way as that in Fig. 4. The only difference is that two sets of pressure data are available. This implies that two sets of sensitivities must be computed (see Eq. B-4) because the equation for the Hessian matrix ($\mathbf{H}_2$ in Eq. A-2) is extended by an additional term.

Again, the optimal test flow may be found by visual inspection. Fig. 6 shows that production should be terminated after only 7 to 25% of the total testing time. This result did not change significantly when other permeability values, within the same range as the above values, were chosen. Closer analysis of the results reveals why. The transient pressure response is almost equally influenced by variations in some of the different elements of Θ. This

TABLE 1—CHARACTERISTIC RESERVOIR PROPERTIES

r_w, m	0.15
r_e, m	750.0
h, m	75.0
ϕ, %	25
μ, Pa·s	1×10^{-3}
c, Pa^{-1}	1×10^{-9}
p_i, MPa	30
V_p, m^3	33.13×10^6
S_w, %	20



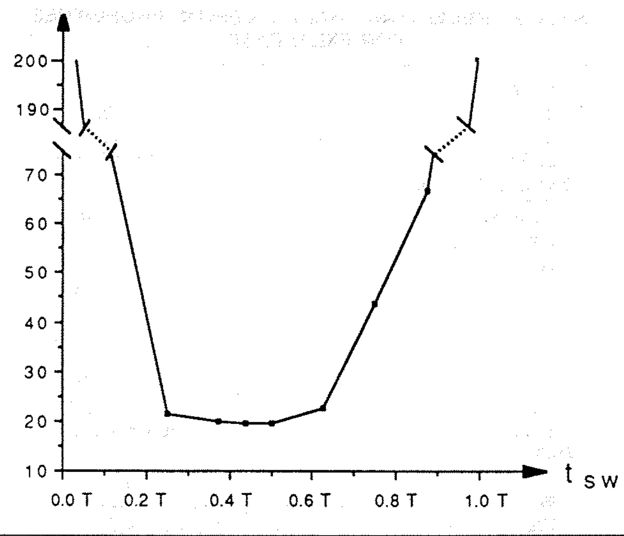


Fig. 4—Value of optimal input design criterion ($F_W=0$, see Eq. A-9) with $\Theta^T=(k,k_z\alpha\beta)$ for different durations of the flow period (see Fig. 3).

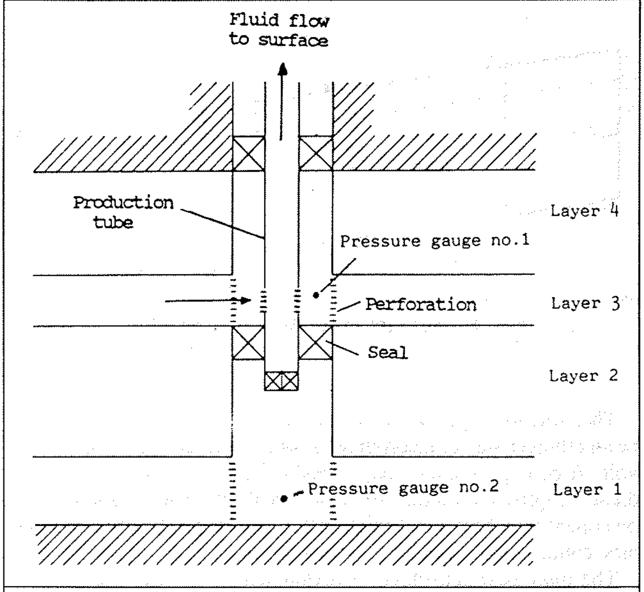


Fig. 5—Schematic of production and monitoring equipment.

is seen by simulation of the sensitivity equations (Eqs. B-3 and B-4). Hence, the transient correlation as defined in Eq. B-5 is high. When the flow period is short, this correlation decreases for the case in question, which increases our ability to compute the parameters (Θ) from the test pressure data with high accuracy.

Example 3: Optimizing the Simultaneous Production Sequence From Multiple Wells

Example 3 is more complicated. **Figs. 7 and 8** depict the reservoir geometry, and **Table 2** gives the reservoir's characteristic properties. The single-phase model is discretized according to a Cartesian grid. An implicit integration algorithm is used. The model consists of 1,280 gridblocks, and the well pressures are computed with Peaceman's[3] equation by assuming that each well is in the center of a gridblock. The unknown parameters (permeabilities and porosities) are assumed to be space-dependent according to the following functions:

$$k_x=k_y=a_1 i+a_2 j+b, \quad (11)$$

$$\phi=a_3(a_1 i+a_2 j+b), \quad (12)$$

$$k_{z1}=a_4(a_1 i+a_2 j+b), \quad (13)$$

$$k_{z2}=a_5(a_1 i+a_2 j+b), \quad (14)$$

and $k_{z3}=a_6(a_1 i+a_2 j+b)$, (15)

where $\Theta^T=[a_1 a_2 a_3 a_4 a_5 a_6]$, k_x,k_y=horizontal permeabilities, i,j=horizontal coordinates (Fig. 7), and b is known. This parameterization assumes that (1) no directional permeability exists in the x-y plane, (2) porosity and horizontal and vertical permeability are correlated, and (3) porosity and permeabilities vary linearly with respect to the spatial coordinates.

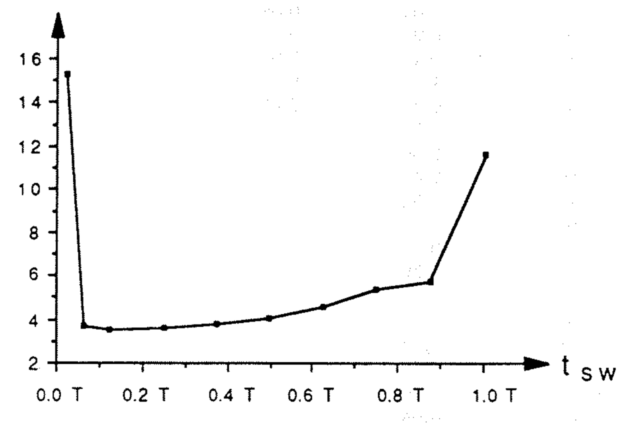
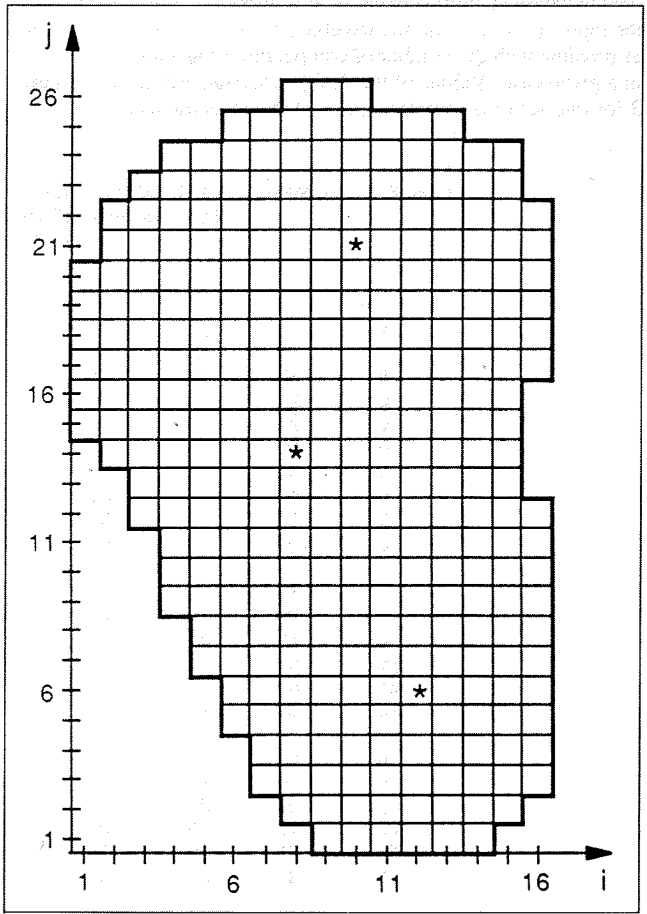


Fig. 6—Value of optimal input design criterion ($F_W=1.0\times 10^{-5}$, see Eq. A-9) with $\Theta^T=(k, k, k, k, k, \beta)$ for differ-



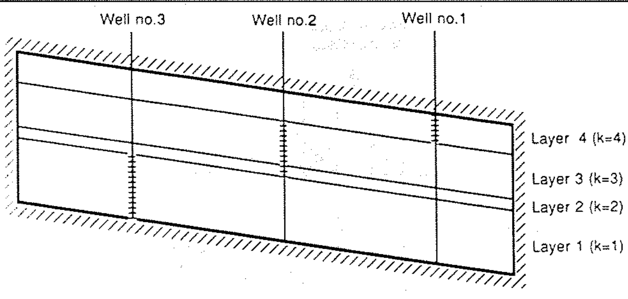


Fig. 8—Vertical reservoir cross section.

The functional parameterization used here can, in some cases, be an efficient way to incorporate prior knowledge regarding a reservoir. A geological study may reveal a depositional process that indicates a certain structure in the permeability distribution that can be copied by a functional relationship, leaving only a few parameters within such a function as unknowns.

The input design problem is formulated by monitoring the pressures and flow rates from the three different wells for 12 days at a sampling time of 2 hours and 45 minutes. Start and end times of production periods from each well are chosen freely. Then, production rates are fixed at 175 m^3/d.

The degree of freedom associated with the choice of the optimal input sequence is increased substantially because there are six variables (i.e., the start and end time instants of flow from each well) that determine the input sequence as opposed to the one variable (t_{sw}) in the well-test case.

The design problem outlined above will have computational limitations because of the simulator's high dimensionality and the need to compute a large number of sensitivities (Eqs. B-1 through B-4). In Example 3, the design criterion (g in Eq. 6) can be computed only in a limited number of cases. It is important, however, to cover the input space (i.e., all the feasible production sequences) as well as possible with the number of computations that can be performed in a given case. Values of the design criterion are shown in **Table 3** for one set of parameter values (denoted as the base case). Permeability and porosity values for selected gridblocks in this case (Figs. 7 and 8) are given in **Table 4**.

TABLE 2—RESERVOIR CHARACTERISTIC PROPERTIES FOR FIELD CASE

r_w, m	0.15
h, m	30.0
μ, Pa·s	1×10^{-3}
c, Pa^{-1}	1×10^{-9}
Discretization, m	
$\Delta x=\Delta y$	100.0
Δz_1*	12.0
Δz_2	3.0
Δz_3	9.0
Δz_4	6.0
z_1**	6.0
z_2	13.5
z_3	19.5
z_4	27.0
Gridblocks	$320\times4=1{,}280$
Wells	3
Well 1	$\beta_1=\beta_2=\beta_3=\infty$, $\beta_4=1.0$†
Well 2	$\beta_1=\infty$, $\beta_2=\beta_3=1.0$, $\beta_4=\infty$
Well 3	$\beta_1=1.0$, $\beta_2=\beta_3=\beta_4=\infty$

*Δz is layer thickness.
**z is a vertical gridpoint.
†$\beta=\infty$ denotes no perforation; see Fig. 8.

Table 3 shows that several production sequences exhibit similar values for the design criterion. Hence, although the degree of freedom for specifying the production sequence has increased substantially compared with the well-test case, the search for a "good" input sequence does not need to be complicated to the same degree.

One major problem is that, again, the design criterion varies as the parameter values change. For some of the results in Table 3, the Hessian matrix was computed for the case where the interlayer permeabilities, k_z, were divided by 10. These results show that the relative difference between "good" and "bad" production sequences are similar, but the ranking of the input sequences is quite different. Therefore, it is important to run the design procedure for different parameter values within the probable range of values.

TABLE 3—COMPUTATIONS OF THE INPUT DESIGN PERFORMANCE CRITERION (Eq. A-9) FOR DIFFERENT PRODUCTION SEQUENCES

Input Sequence*						Value of Design Criterion ($F_W=1.0\times10^{-5}$)	
t_{10}	t_1	t_{20}	t_2	t_{30}	t_3	Base Case	$k_z/10$
0	25	0	25	0	25	1.67	0.44
0	25	0	25	0	50	0.99	0.37
0	25	0	50	0	25	0.89	0.34
0	25	0	50	0	50	0.99	0.32
0	50	0	25	0	25	0.53	0.59
0	50	0	25	0	50	0.98	0.75
0	50	0	50	0	25	0.63	0.55
0	50	0	50	0	50	1.79	0.61
0	100	0	100	0	100	6.45	2.80
0	25	0	50	25	75	0.54	
0	25	25	75	0	50	0.40	
0	25	25	75	25	75	0.38	
25	50	0	50	0	50	1.15	
25	50	0	50	25	75	0.94	
25	50	25	75	0	50	0.79	
0	50	0	25	25	50	0.84	
0	50	25	50	0	25	0.48	
0	50	25	50	25	50	0.92	
25	75	0	25	0	25	0.46	
25	75	0	25	25	50	0.64	
25	75	25	50	0	25	0.54	

Guidelines for Practical Use: Single-Well-Test Case

Guideline 1. Decide the degrees of freedom needed to specify the flow-rate sequence. For instance, in Examples 1 and 2, there was only one free variable—the length of the production period. If the total test time is also a free variable, however, then the number of free variables equals two.

Guideline 2. Extend a well simulator by computing the Hessian matrix as shown in Eqs. A-2, A-3, B-3, and B-4. If several pressure gauges measure different pressures, as in Example 2, include additional terms in the $\mathbf{H}_2$ computation of Eq. A-2.

Guideline 3. Include the computation for the input design criterion (Eq. A-9). The weighting factor, F_W, may initially be chosen to equal zero. It is a tuning parameter[2] and may be adjusted from experience.

Guideline 4. Include a graphic representation similar to Fig. 4 or 6 that shows the input design criterion as a function of the design variables. In this case of more than one free variable, use multiple graphs.

Guideline 5. Choose the simulator parameters, Θ, that are to be estimated by the well-test data.

Guideline 6. Compute the design criterion for different Θ parameter values (within their probable range of values). The range is based on information from well-log and seismic data and on geological knowledge. Finally, choose by visual inspection the flow sequence that, on average, minimizes the design criterion.

Making the necessary software extensions to a standard simulator (Guidelines 2 through 4) gives a design procedure consisting of Guidelines 1, 5, and 6.

Discussion

Currently, no *generic* methods are available for designing flow sequences that use the information content of the data as design parameters. Certain rules of thumb exist in well testing. In the North Sea, for instance, a well test is typically run so that the buildup period is twice as long as the drawdown period. In a multiple-well case, such as Example 3, there are no comparable methods.

The advantage of the proposed method is relatively easy to demonstrate. In a design, the specifications from a conventional design method should be compared with the specifications for the proposed method. The certainty by which the Θ parameters may be computed for a given design can be calculated (see Eq. A-7). Hence, it is possible to perform a quantitative benefit analysis like those in Examples 1 through 3 and examples given by Foss.[4] These examples show that, on average, the maximum variance of the parameter estimates decreases by 50% with active use of the proposed method. Note that the proposed method also improves the conditioning of the parameter space. In practice, this increases the chances of converging to sensible parameter estimates in the parameter estimation run for the data obtained.

The proposed method is an off-line method in the sense that the design is performed before a test. The method can also be used as a design tool while a test is run. Assume that a flow sequence has been specified and that the free variable is the length of the flow period. The following procedure may now be chosen.

Step 1. Estimate Θ parameters at, for instance, every $0.1T$ (T=total test period) for the data collected up to this time. The rationale is that the parameters may be chosen with increased confidence after the test is run for some time rather than before the test.

Step 2. Perform a new design of the flow-rate sequence for the estimated Θ parameter values from Step 1.

Step 3. If the switch-off time, t, given by the redesign is not within the future time interval $0.1T$, the procedure is repeated after the test has run another $0.1T$. In the opposite case, the production is terminated according to the design specification (the pressure is monitored until $t=T$).

This approach will improve the design. This procedure, however, does assume that good-quality data are available in the early parts of a drawdown period. The design system can be implemented in a standard computer system provided that it can withstand

TABLE 4—PERMEABILITY AND POROSITY VALUES FOR SELECTED GRIDBLOCKS IN EXAMPLE 3 (Figs. 7 and 8)

Coordinate		Layer 1		
i	j	k_x,k_y (md)	k_z (md)	ϕ (%)
12	6	400	12	20
8	14	520	16	20
10	21	680	20	34

A shortcoming of the proposed method is that it is based on a linear analysis because nonlinear models must be linearized to compute the Hessian matrix. This poses no major problem in the case of mild nonlinearities. Severe nonlinearities (as in multiphase models), however, do complicate matters because a first-order linearization is valid only in a limited region about the linearized values. Hence, frequent linearizations are necessary, and the computational requirements must be increased substantially. To assess the input design method in these cases, trials must be run with nonlinear models.

Modeling errors represent a serious problem that can disrupt any analysis of the kind discussed in this paper. The reason is that, in such a case, the analysis is performed on a model that describes a system different from the reservoir in question. Two measures should be included to solve this problem. First, when the model is being chosen, it is essential to use as much of the knowledge available before the experiment as possible. Second, the design calculations should be performed with different parameter values within their probable range of values. In some cases, when the structure of the model itself is doubtful (e.g., because of uncertain layering or the possible presence of discontinuities in the layers), different models should be used.

Up to now, the discussion has focused mainly on the single-well-test case. Two applications of the design method, however, are of particular interest in the field case. First, the design of early production (interference) tests—i.e., tests run during the period before normal production—are expensive and make the design phase especially important. Second, the method may be used to specify the flow-rate variations around the plateau rates during normal production to improve the information content in the data used for history matching.

Because the design method is generic, it is applicable to other types of design problems, such as the choice of time-varying inputs in laboratory experiments and the selection of perforation intervals in well testing. Foss[4] investigated this.

Conclusions

1. The optimal input design method is applicable to design problems in reservoir evaluation and well testing.
2. The proposed method markedly increases the information content in the data; hence, the uncertainty by which the parameters may be computed from the measured data decreases with active use of the proposed method.

Nomenclature

a = unknown parameter
$\mathbf{A(k)}$ = tridiagonal matrix
b = known scalar
c = compressibility, Pa^{-1}
C_{ma} = covariance matrix
e = random (white) noise on data
f = function
F_c = correlation factor
F_W = weighting factor in design criterion
g = functional defining input design criterion
h = reservoir thickness, m
$\mathbf{H}$ = Hessian matrix

k_r, k_z = permeability in 2D cylindrical geometry, md
k_s = permeability in damaged zone, md
k_x, k_y = permeability in 2D Cartesian geometry, md
M = scalar number
N = total number of data points
p = pressure, Pa
p_w = well pressure, Pa
p_{12} = pressure in gridblock adjoining perforation, Pa
Δp = pressure drop into the wellbore, Pa
q = input; production flow rate, m^3/d
q_s = sandface flow rate, m^3/d
Q = set of admissible inputs; production sequence
r = radial axis, m
r_e = external radius, m
r_w = wellbore radius, m
R = ID Euclidean space
R^M = M-dimensional Euclidean space
S = wellbore storage constant
S_w = water saturation, percent
t = time, seconds
t_{sw} = stabilizing time of flow period in a well test
T = end time or total testing time
V = eigenvector matrix
V_p = pore volume, m^3/d
x = Cartesian coordinate
y = computed output; Cartesian coordinate
y_{OB} = measured/observed output data
z = vertical axis
α = parameter associated with wellbore storage constant
β = parameter associated with pressure drop past damaged zone
ϵ = error between observed and computed output
$\mathcal{E}[\ ]$ = expectational operator
Θ = unknown parameter
$\bar{\Theta}$ = error in the computed value of Θ
κ = condition number of a matrix
λ = eigenvalue
Λ = eigenvalue matrix
μ = viscosity, Pa·s
ϕ = porosity, percent

Subscripts

i,j = horizontal gridblock coordinates
max = maximum
min = minimum

Superscripts

n = time index
T = transpose
$\rightarrow$ = vector

Acknowledgment

I thank Jon Kleppe of the Dept. of Petroleum Engineering and Applied Geophysics at the U. of Trondheim for his helpful comments and suggestions during the preparation of this paper.

References

1. Dogru, A.H. and Seinfeld, J.H.: "Design of Well Tests to Determine the Properties of Stratified Reservoirs," paper SPE 7694 presented at the 1979 SPE Reservoir Simulation Symposium, Denver, Jan. 31-Feb. 2.
2. Foss, B.A.: "A Modified Optimal Input Design Criterion," *Proc.*, IEEE Conference on Decision and Control, Los Angeles (Dec. 9-11, 1987) 303-08.
3. Peaceman, D.W.: "Interpretation of Well-Block Pressures in Numerical Reservoir Simulation," *SPEJ* (June 1978) 183-94; *Trans.*, AIME, **265**.

6. Mehra, R.M.: "Optimal Input Signals for Parameter Estimation in Dynamic Systems—Survey and New Results," *IEEE Automatic Control, Trans.* (1974) 753-68.

Appendix A—Optimal Input Design Method

An expression for parameter uncertainty can be derived by considering the parameter estimates as stochastic variables. Assume that the parameter estimation is based on minimizing the following loss function; a quadratic criterion is used as an example.

$$J(\Theta)=\frac{1}{2}\sum_{n=1}^{N}[y_{OB}^{n}-y(\Theta)^{n}]^2=\frac{1}{2}\sum_{n=1}^{N}(\epsilon^n)^2,\ \dim(\Theta)=M, \quad \text{(A-1)}$$

where y_{OB}^n and y^n are the observed and computed outputs, respectively (for instance, by use of a simulator), at time n. The simulated output depends on the unknown Θ parameters.

If J is twice differentiable with respect to Θ, the Hessian matrix can be computed by

$$\mathbf{H}(\Theta)=\frac{\partial^2 J}{\partial\Theta\partial\Theta^T}=\mathbf{H}_1+\mathbf{H}_2=\frac{1}{N}\sum_{n=1}^{N}\left(\frac{\partial^2\epsilon^n}{\partial\Theta\partial\Theta^T}\right)\epsilon^n$$

$$+\frac{1}{N}\sum_{n=1}^{N}\frac{\partial y^n}{\partial\Theta}\frac{\partial y^n}{\partial\Theta^T}, \quad \text{(A-2)}$$

where, in the interest of simplicity, it is assumed that y is a scalar.

If Θ_0 is assumed to be the optimal parameters for which an ideal match is obtained [i.e., the model is an exact replica of the physical system in question, $J(\Theta_0)=0$], then the implication is that

$$\mathbf{H}(\Theta_0)=\mathbf{H}_2(\Theta_0) \quad \text{(A-3)}$$

because all $\epsilon^n=0$ in the case of no noise on the observed data.

If, on the other hand, there is noise on the observed data, and this noise is additive and white,

$$y_{OB}^n=y(\Theta_0)^n+e^n, \quad \text{(A-4)}$$

and the system that generates y is linear, it is straightforward to show that the following equation holds (see Ref. 4).

$$\mathcal{E}[\bar{\Theta}^T\mathbf{H}(\Theta_0)^{-1}\bar{\Theta}]=\text{var}(e),\ \bar{\Theta}=\Theta-\Theta_0, \quad \text{(A-5)}$$

where $\mathbf{H}$ is a positive semidefinite matrix. It can be eigenvalue-decomposed by

$$\mathbf{H}=V\Lambda V^T, \quad \text{(A-6)}$$

where $\Lambda=\text{diag}\{\lambda_1\ldots\lambda_M\}$ and $V=\{V_{ij}\}$. An approximate expression of the variance of one of the estimates can be given by

$$\text{var}(\bar{\Theta}_i)=\max_j\left[\frac{V_{ij}^2}{\lambda_j M}\text{var}(e)\right]. \quad \text{(A-7)}$$

It is common in optimal input design theory to assume that an unbiased and minimum variance estimator is used; hence, it is just the variance that determines the uncertainty. The minimum variance (efficient) estimate can be calculated by the use of the Fisher information matrix.[5] Under suitable conditions,[6] the Fisher information matrix can be approximated with the Hessian matrix in Eq. A-3. The input design method may now be written in a compact form,

$$\min_{q\in Q}\ g(\mathbf{H}^{-1}), \quad \text{(A-8)}$$

by recognizing that $\mathbf{H}$ directly influences the uncertainty of the parameter estimates and that the input, q, to the system in question influences $\mathbf{H}$. Foss[2] and Mehra[6] propose several choices. The one used in this investigation is

$$g(\mathbf{H}^{-1})=\lambda_{\max}(\mathbf{H}^{-1})+F_W\kappa(\mathbf{H}),\ 0\le F_W\le\lambda_{\min}(\mathbf{H}^{-1}). \quad \text{(A-9)}$$

The first term places emphasis on reducing parameter uncertainty keep the condition number as low as possible to improve the robustness of the parameter estimation runs that follow the experiment. When $\lambda_{max}(\mathbf{H}^{-1})$ shows a flat minimum, it is especially advantageous to incorporate the condition number in the criterion.[2]

Appendix B—Computing the Hessian Matrix

The Hessian matrix is computed with Eqs. A-2 and A-3; hence, an expression of the sensitivities $\partial y^n/\partial\Theta$ is needed. For example, consider that the reservoir is described by the following implicit linear, discrete model:

$$\mathbf{A}(\Theta)p^{n+1}=p^n+\mathbf{B}(\Theta)q^n \quad \text{(B-1)}$$

$$\text{and } y^n=\mathbf{D}(\Theta)p^n+\mathbf{E}(\Theta)q^n, \quad \text{(B-2)}$$

where q is the production term. Eq. B-2 is the observation equation. Equations for the sensitivity with respect to one of the parameters Θ_i within Θ are derived by differentiating the above equation with respect to Θ_i.

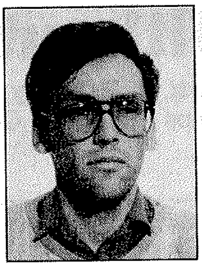
$$\mathbf{A}(\Theta)\frac{\partial p^{n+1}}{\partial\Theta_i}=\frac{\partial p^n}{\partial\Theta_i}+\frac{\partial\mathbf{B}(\Theta)}{\partial\Theta_i}q^n-\frac{\partial\mathbf{A}(\Theta)}{\partial\Theta_i}p^{n+1},\ i=1\ldots M, \quad \text{(B-3a)}$$

$$\frac{\partial p^1}{\partial\Theta_i}=0,\ \ i=1\ldots M, \quad \text{(B-3b)}$$

$$\text{and } \frac{\partial y^n}{\partial\Theta_i}=\mathbf{D}(\Theta)\frac{\partial p^n}{\partial\Theta_i}+\frac{\partial\mathbf{D}(\Theta)}{\partial\Theta_i}p^n+\frac{\partial\mathbf{E}(\Theta)}{\partial\Theta_i}q^n,\ i=1\ldots M. \quad \text{(B-4)}$$

The initial conditions that are equal to zero stem from the assumption that the initial pressure is not affected by Θ. For a nonlinear model, the sensitivity equations are computed by linearizing the model equations about their operating point.

Author



Bjarne A. Foss is a professor of engineering cybernetics at the Norwegian Inst. of Technology, U. of Trondheim. His interests are in applying cybernetic methods to reservoir evaluation to improve the quality and interpretation of reservoir data. He previously worked with SINTEF, a Norwegian research company, applying advanced control concepts in the Norwegian industry. Foss holds MS and Dr. ing. degrees from the U. of Trondheim.

The correlation factor between two parameters Θ_i and Θ_j is defined by

$$F_c(\Theta_i,\Theta_j)=\mathbf{H}_{ij}/\sqrt{\mathbf{H}_{ii}\mathbf{H}_{jj}}\ ,\ F_c(\Theta_i,\Theta_j)=[-1,1],\ i\neq j. \quad \text{(B-5)}$$

This is a measure of the dynamic correlation between the sensitivities of the two parameters Θ_i and Θ_j. $F_c=\pm1$ implies that a change in either Θ_i or Θ_j gives the same transient change in the output y. Hence, the two parameters are not distinguishable. In general, low dynamic correlation improves parameter identifiability; i.e., the error bounds with which parameters may be estimated are lowered.

SI Metric Conversion Factors

bbl	×	1.589 873	E−01	=	m^3
cp	×	1.0*	E−03	=	Pa·s
ft	×	3.048*	E−01	=	m
md	×	9.869 233	E−04	=	μm^2
psi	×	6.894 757	E−03	=	Pa
psi^{-1}	×	1.450 377	E−04	=	Pa^{-1}

*Conversion factor is exact.

SPEFE

Original SPE manuscript received for review April 18, 1988. Paper (SPE 18402) accepted for publication March 28, 1990. Revised manuscript received March 9, 1990.