Input signal design to estimate interwell connectivities in mature fields from the capacitance-resistance model

Gustavo A. Moreno[1, 2, 3]* and Larry W. Lake[3]

[1] YPF-Tecnología, Baradero S/N (1925) Ensenada, Buenos Aires, Argentina

[2] YPF, Macacha Güemes 515, CABA, Argentina

[3] CPARM, University of Texas at Austin, Austin, TX, USA

Abstract: Interwell connectivities are fundamental parameters required to manage waterfloods in oil reservoirs. Data-driven models, such as the capacitance-resistance model (CRM), are fast tools to estimate these parameters from time-correlations of input (injection rates) and output (production rates) signals. Noise and structure of the input time-series impose limits on the information that can be extracted from a given data-set. This work uses the CRM to study general prescriptions for the design of input signals that enhance the information content of injection/production data in the estimation of well-to-well interactions. Numerical schemes and general features of the optimal input signal strategy are derived for this problem.

key words: Intervell connectivity, capacitance-resistance model, mature field, experiment design

1 Introduction

Data driven models are an alternative to reservoir simulation, with the aim of giving an over-determined description of the reservoir response to an external stimuli. The capacitance-resistance model (CRM) (Yousef et al, 2006) is a fast tool that allows the estimation of interwell connectivities based on time-correlations of input (injection rates) and output (production rates) signals via discrete mass balance equations. Its importance relies on the fact that this is the minimal (low order) model to infer interwell connectivities from dynamic field data. Data analysis based on this scheme has been successfully applied in many waterflood studies (Sayarpour et al, 2009; Weber, 2009; Nguyen et al, 2011).

However field-data is affected by noise and thus well-to-well interactions can only be calculated within a limited accuracy that depends both on noise amplitude and time-series structure of the data. The first attempts to quantify the accuracy and conditions to obtain a well-posed problem were done by Albertoni and Lake (2003) and Yousef et al (2006). Recently we showed that errors in unconstrained analytical estimates of interwell connectivity can be used as an upper bound for constrained problems (Moreno and Lake, 2014), even taking into account the non-zero response time in producers. This work uses these results to analyze input

design for mature waterfloods using system identification techniques. We find general prescriptions to optimize the synchronization of injection signals enhancing the information content of the dataset.

The following section reviews the CRM and uncertainty bounds for the constrained optimization problem. In Section 3 we apply input signal design techniques to this problem and give a framework to solve the general case numerically. Section 4 solves approximated versions of the problem semi-analytically and gives new insight in general prescriptions to obtain an optimum scheme. In Section 5 we summarize the results and their implications.

2 Interwell connectivities from the CRM and uncertainty

The CRM is based on a discrete material balance (Yousef et al, 2006) that connects rates from each injector well $I_i(t)$ to the total liquid rate in an associated producer $q_j(t)$. Assuming constant total compressibility, a linear relationship between well pressure and flow, and the constant bottom hole flowing pressures the evolution equation for the system is

$$\tau_j \dot{q}_j(t) + q_j(t) = \sum_\alpha f_{i(j,\alpha)j} I_{i(j,\alpha)}(t) \tag{1}$$

Here the index $\alpha=1, 2, \ldots, n_j$ spans the list of neighbors of the j-th producer, then n_j is the maximum number of neighbors of the j-th producer, and the mapping $i(j,\alpha)$ stands for the i-th injector (of the full list of injectors) that is connected to the j-th producer. This notation may seem cumbersome at first,

---

*Corresponding author. email: gustavo.moreno@ypftecnologia.com, gustavo.moreno@ypf.com

but it will be simpler to span the indexes listing the neighbors per producer because here only neighboring interactions to the j-th producer are important (of course we still do not define a priori the neighbor list). The time constant τ_j is a measure of the response time in the producer and $f_{i(j,\alpha)j}$ are the connectivities to the neighboring injectors. In what follows we will assume that the time constants are known. However this assumption does not yield a loss of generality in practical cases because the model applies only when small time constants are present. Moreover, the connectivity fits are not sensitive to these parameters (Kaviani, 2012).

To obtain the production rate, Eq. (1) can be integrated

$$q_j(t) = q_j(0)e^{-t/\tau_j} + \sum_\alpha f_{i(j,\alpha)j} \left\langle I_{i(j,\alpha)} \right\rangle_j (t) \qquad (2)$$

with

$$\left\langle I_{i(j,\alpha)} \right\rangle_j (t) = \int_0^t \frac{e^{-(t-s)/\tau_j}}{\tau_j} I_{i(j,\alpha)}(s)\,\mathrm{d}s \qquad (3)$$

Provided that the time window is much larger than the typical response time τ_j for almost all times we have $t/\tau_j \gg 1$ and thus:

$$q_j(t) \cong \sum_\alpha f_{i(j,\alpha)j} \left\langle I_{i(j,\alpha)} \right\rangle_j (t) \qquad (4)$$

Then, the production rate is linearly coupled to the averaged (in the sense of Eq. (3)) injection rates.

The values of the connectivities can be obtained by least squares fit of the predicted rates $q_j(t)$ to the monthly measured rates $Q_j(t_k)$. This is a constrained quadratic programming problem that can be solved in a number of ways (Bjorck, 1996; Boyd and Vandenberghe, 2004). The output of the fitting should give information about the reservoir, but it is of course affected by uncertainties on the input data. The uncertainty in the unconstrained problem has been analyzed in different contexts (Ljung, 1999; Weber, 2009), however only recently it has been shown that unconstrained analytical estimates can be used as an upper bound for the error estimation in interwell connectivities (Moreno and Lake, 2014). The associated Fisher information matrix (Ljung, 1999) in the present problem is block-diagonal where each block reads

$$A_{\alpha\beta}^{(j)} = \sum_k \left\langle I_{i(j,\alpha)} \right\rangle_j (t_k) \left\langle I_{i(j,\beta)} \right\rangle_j (t_k) \qquad (5)$$

This is, each block is associated with one producer and the dimension of the block is equal to the number of injection wells that are possibly connected to that producer (for example, but not necessary only, nearest neighbors). In this work we assume that the influencing wells for each producer are known, Moreno and Lake (2014) studied the general problem of information-degradation when this is not available a priori. The deviation of the parameters in the constrained problem from their mean values $\delta f_{i(j,\alpha)j} = f_{i(j,\alpha)j} - \left[f_{i(j,\alpha)j} \right]$ (here[.] denotes average over noise realizations) when the noise is uncorrelated is bounded by

$$\left[\delta f_{i(j,\alpha)j} \right]^2 = A_{\alpha\alpha}^{-1(j)} \sigma_j^2 \qquad (6)$$

where σ_j is the standard deviation of the noise in the j-producer rates. Gaussian uncorrelated errors are not the most general type of model, and some correlated sources may exist in practice. However, the case considered here is the first non-trivial approximation to the problem and addresses widely used assumptions in the literature. Thus the Fisher information matrix (with blocks given by Eq. (5)) can be used to design the input injection in such a way that the uncertainty in the parameter estimation is a minimum. The mathematical problem is to choose an injection scheme $\left\{ I_{i(j,\alpha)}(t) \right\}$ that under some optimality criteria can minimize this uncertainty. In the following sections we analyze this problem in detail.

3 Input design: general case

Input design is a well-known problem in system identification and control communities (Ljung, 1999; Box et al, 2008) where many aspects have been studied (Fedorov, 1997; Zarrop, 1979). In the present context we have seen how the unconstrained case can be used to set an upper bound on the errors in the parameter estimates. Thus future injection signals can be designed to minimize the upper bound. This will be the best possible input, in the sense of information content, to fit the model without any other a priori knowledge of the system. The problem can be treated using techniques for input design. The first task is to select an optimality criterion which will give the relative weights to the uncertainty growth factors $A_{\alpha\alpha}^{-1(j)}$. There are many possible optimality criteria (Boyd and Vandenberghe, 2004), here, to treat the general case, we chose the so-called A-optimal design, with objective function:

$$S = \sum_{j,\alpha} A_{\alpha\alpha}^{-1(j)} \qquad (7)$$

This objective is proportional to the sum of eigenvalues of the uncertainty ellipsoid and results in a well behaved quantity for large systems. It is also the sum of squares of the deviation growth rate bounds for all the injector-producer pairs.

The problem involves an injection rate for each future time period, we assume without loss of generality that the signal $I_{i(j,\alpha)}(t)$ is defined for a certain sequence $t_1, ..., t_k, ..., t_w$. The solution has to take into account the minimum and maximum allowed rates per injector $I_{i(j,\alpha)}{}^{\mathrm{Min}} < I_{i(j,\alpha)}(t_k) < I_{i(j,\alpha)}{}^{\mathrm{Max}}$ and also a similar constraint for the total rate. The latter will guarantee that at least a total rate of I_{min} is being injected and that the total injection limit (typically plant capacity) I_{plant} is not exceeded. These constraints are important to take into account operational restrictions that avoid the interference of the inter-well interaction test with expected total production. The optimization of the input to enhance the informative content of the data is carried out within the gap allowed by productivity constraints.

If the injection rates are discretized the problem can be written in standard form using basis functions (Boyd and Vandenberghe, 2004) but this results in a difficult combinatorial problem. The constraints on the input signal and structure of the model resembles that of input optimization for aircraft parameter estimation (Reid, 1972;

Morelli, 1990); thus a time domain approach is the most adequate to incorporate the restrictions on the input. Not all the techniques from that field can be used straight forwardly here because further constraints have to be included on the injection history. In oil field operation a selective change of injection rates will have a cost, thus Walsh-like (Morelli, 1990) linear combinations of input signals are not convenient. To solve the problem in the general case via a fast implementation we have used high-level commercial software, namely General Algebraic Modeling System (GAMS). Many optimization solvers are available under GAMS, we have selected CONOPT (a general propose non-linear programming solver) to minimize our objective function. In our problem the variables are injection rates for each time period $I_{i(j,\alpha)}(t_k)$ thus, to obtain a smooth solution, we solve different steps that progressively penalize functions with the same performance (in the sense of uncertainty) but with more fluctuations in the rate. Specifically we have included a cost function that measures the total number of changes in the input for a given scale ε, namely

$$C_\epsilon = \sum_{j,\alpha,k\geq1} \frac{\left(I_{i(j,\alpha)}(t_k) - I_{i(j,\alpha)}(t_{k-1})\right)^2}{\epsilon^2 + \left(I_{i(j,\alpha)}(t_k) - I_{i(j,\alpha)}(t_{k-1})\right)^2} \qquad (8)$$

and set the upper bound

$$C_\epsilon \leq C \qquad (9)$$

The constraint on C_ε rule out all functions that are highly fluctuating (in the scale of ε). Moreover, in the limit ε to 0, C_ε counts exactly the total amount of changes in the input signal, favoring piecewise constant signals. Thus, solving the nonlinear programming (NLP) problem for a sequence of geometrically decreasing values of ε and using each solution as initial guess for the following iteration one can find an optimal (or near optimal) solution in typically 6 steps. Fig. 1 shows a simple example where four injectors affect one producer and the total amount of water available to inject coincides with the maximum rate for each single well. As it can be seen, the optimum corresponds to a signal with maximum contrast between the four stimuli on the producer. In the following section we develop analytical tools that can be applied to study this configuration obtaining the same result. When more complex situations that involve many injectors and producers are considered the use of the program becomes nontrivial, because of its ability to synchronize the stimuli in the field within the injection constraints. Fig. 2 shows an example of such case, solving for the optimum input with 9 injectors and 4 producers in an inverted five-spot layout (and also including constraints). This type of problem can be of particular interest for example in pilot design, where the interaction between wells is important to determine the performance of the process being tested. Moreover, for any arbitrary given set of data both the actual and the optimal signals can be compared using the Fisher information matrix, the latter being the ultimate theoretical maximum for the informative content of the set.

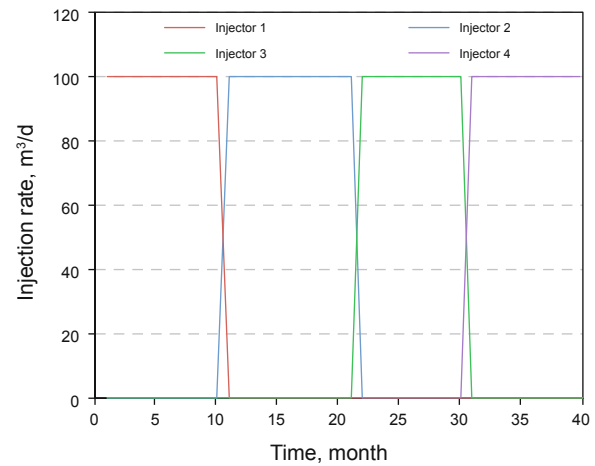The initial step of the program, which here is generated

Fig. 1 Optimal injection scheme to minimize future uncertainty in a 5-spot pattern with the total water injection rate equal to the maximum bound on each injector (100 m³/d). Here the minimum injection rate is zero.

using randomly located Gaussian-like inputs, is not trivial and some difficulties, such as suboptimal solutions, can occur. Nevertheless below we exemplify how these difficulties can be overcome using other approximations.

4 Input design: general features

When the constraints $I_{\min} \leq \sum_i I_i(t_k) \leq I_{\text{plant}}$ and $C_\varepsilon \leq C$ are removed it is possible to use the calculus of variations (Elsgolts, 1980) to show that the optimum strategies for this problem are of the bang-bang type. More generally, as a consequence of the Pontryagin maximum principle, this property of the solution was already remarked upon by Reid (1972) and Morelli (1990). In the general case where all the constraints are included we also find similar results (even releasing the constraint $C_\varepsilon \leq C$ which allows for general profiles). However we provide no general proof here; in what follows we argue why this kind of solution is expected (small cases, such as two injectors with one producer, can be solved exactly by exhaustive search in the low response time limit finding that this is valid).

To provide further insight in the optimum strategy (general features of the synchronization, shape of the input signal, and even GAMS-free algorithms) we will further assume that the time constants τ_j are small compared to the time-step. This does not imply a loss in generality because high frequency components in the input signal are linearly suppressed (as it can be seen for a Fourier decomposition of Eq. (3)), thus components with frequency much higher than $1/\tau_j$ are washed out, and the signal can be approximated by a piecewise constant signal (one value per time-step) that represents components of frequency lower than $1/\tau_j$.

In such case the Fisher information matrix for a sequence of W time periods $t_1, \ldots, t_k, \ldots, t_w$ let's call it M_W, is simply the sum of W matrices formed with the local value of the injection rates interacting with the n producers:
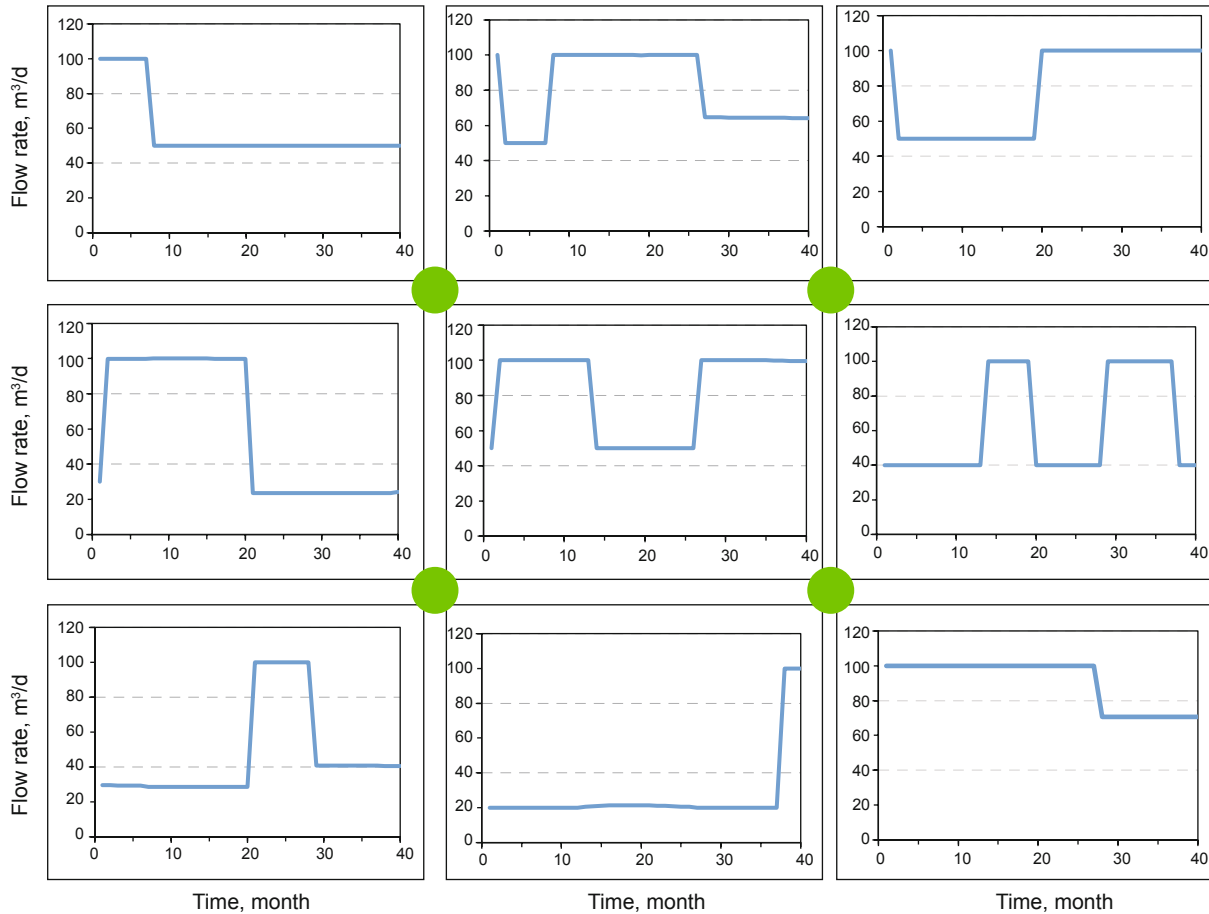
$$M_W = \sum_k m(t_k) \qquad (10)$$

Fig. 2 Optimal injection scheme (rate (m³/d) vs. time (month)) to minimize future uncertainty in a block of four inverted 5-spots. Here injectors have non-zero minimum rates and their *XY*-position in the field is represented by their injection-rate plot. The optimal scheme results in a synchronized bang-bang type input ("bang-bang" means that the signals change between the upper and lower imposed limits). Producers (green dots) are connected only to nearest neighbors (injectors represented by their injection-rate plots).

with

$$m(t_k) = v(t_k)v(t_k)^t \tag{11}$$

and

$$v(t_k) = \begin{bmatrix} I_{i(1,1)}(t_k) \\ \vdots \\ I_{i(j,\alpha)}(t_k) \\ \vdots \\ I_{i(n,\alpha_n)}(t_k) \end{bmatrix} \tag{12}$$

Suppose M_W is given and that we want to select $v(t_{W+1})$, that is the vector of injection rates for the following month, in such a way that the choice minimizes uncertainty for M_{W+1}. That will be the case for example if we want to solve the optimum strategy for the following month given a certain injection history. To solve this problem we can use E-optimality criteria (Zarrop, 1979) (in the general case this criterion makes the problem difficult, but in this case it can be handled). This means that we have to minimize the largest eigenvalue of M_{W+1}^{-1} or, equivalently, maximize the smallest eigenvalue of M_{W+1}. So let $v_l^{(W)}$ be the unit-norm eigenvector

associated to the smallest eigenvalue $\lambda_l^{(W)}$ of M_W (if the subspace is degenerate the same argument can be applied to each element of the basis), if we set $v(t_{W+1})=\mu v_l^{(W)}$ we have

$$M_{W+1}v_l^{(W)} = \left(\lambda_l^{(W)} + \mu^2\right)v_l^{(W)} \tag{13}$$

Meanwhile for the remaining elements of the basis

$$M_{W+1}v_r^{(W)} = \lambda_r^{(W)}v_r^{(W)} \forall r \neq l \tag{14}$$

that is, all the other eigenvectors (assumed without loss of generality orthogonal to $v_l^{(W)}$) still remain being eigenvectors of the same eigenvalue. Thus, clearly, this choice maximizes the contribution of the next measurement to the increase in the lowest eigenvalue of M_W, which is the same as saying that this choice minimizes uncertainty. However, because of the constraints on $v(t_{W+1})$ we will not always be able to select a vector contained only in the subspace generated by $v_l^{(W)}$. Using the latter observations in the general case, we will search for a vector $v(t_{W+1})$ with the largest projection onto $v_l^{(W)}$ satisfying the constraints $v(t_{W+1})$ in U (U being the feasible region). This guess can be explicitly demonstrated using perturbation theory, because the smallest eigenvalue $\lambda_l^{(W+1)}$ of M_{W+1} (to leading order in the perturbation) is approximately given by:

$$\lambda_l^{(W+1)} = \lambda_l^{(W)} + \left(v\left(t_{W+1}\right), v_l^{(W)} \right)^2 \qquad (15)$$

Thus the best choice to maximize the smallest eigenvalue is the election of an injection rate in the feasible region with maximum projection onto $v_l^{(W)}$. The maximum projection problem under the constraints can be solved via two linear programming problems:

P1: Minimize $v_1\left(t_{W+1}\right)v_l^{(W)}$, $v_1\left(t_{W+1}\right)\epsilon U$

P2: Maximize $v_2\left(t_{W+1}\right)v_l^{(W)}$, $v_2\left(t_{W+1}\right)\epsilon U$

because the feasible region U is described by linear inequalities. Finally select $v(t_{W+1})$ from the maximum of the two options $\{(v_1(t_{W+1}), v_l^{(W)})^2, (v_2(t_{W+1}), v_l^{(W)})^2\}$. This method can be applied for input signal design using only linear programming and the smallest eigenvector/eigenvalue determination, which are easily and efficiently implemented in any programming language. The latter also indicates that bang-bang type inputs should be expected here because the linear programming problems will find their optimum at the boundaries of the feasible region.

Let us review the case of the inverted five-spot with 4 injectors as an extension of the application of the previous ideas. In this example each injection rate is in the interval [0, I_0] and the total rate is I_0. The injection rate for each element can be represented geometrically as one dimension in R^4, the problem is then to select a sequence of points in R^4 that represents the piecewise constant values of the optimum input signal. With no previous history, the minimum eigenvalue of the information matrix is zero for all vectors, we thus have full degeneracy. We start applying a stimulus to an arbitrary

direction for k months, as shown in Fig. 3 for I_1. After this we have to distribute water among the injectors again. To minimize uncertainty the new stimuli must lay on the orthogonal space spanned by injectors $\{I_2, I_3, I_4\}$ (strictly this is a bounded set, not a subspace). In principle any direction in such set will be equivalent, in the figure we applied k months of injection rate I_0 to injector I_2. The following step involves now a 2D set which can be treated in a similar way. The final result is matrix proportional to the identity (proportionality factor equals kI_0^2) that minimizes the overlap of different stimuli maximizing the information of the dataset. The signals for each step are taken to be orthogonal and on one of the principal directions $\{I_1, I_2, I_3, I_4\}$, this is the best alternative for global optimization because rates cannot be negative and thus the sequence of orthogonal spaces is restricted, principal directions maximize orthogonal projections inside the feasible region. Despite the fact that this simple case with 4 injectors can be tackled by hand, it should be emphasized that the previous algorithm does not apply for global optimization but only to select the following step injection distribution. Nonetheless this method is applicable to many situations of interest. Suppose for example that a given dataset is going to be enlarged with new measurements in the field, and that one has generated an estimation of the time constants for each producer (now not necessarily small). The question is then how to select the injection rates for the following months in such way that they are within operational constraints and also give as much information as possible. In this case we can construct the Fisher information matrix of the input signal (Eq. (5))with the available history and find its minimum eigenvalue. This eigenvalue, in the general case, will be
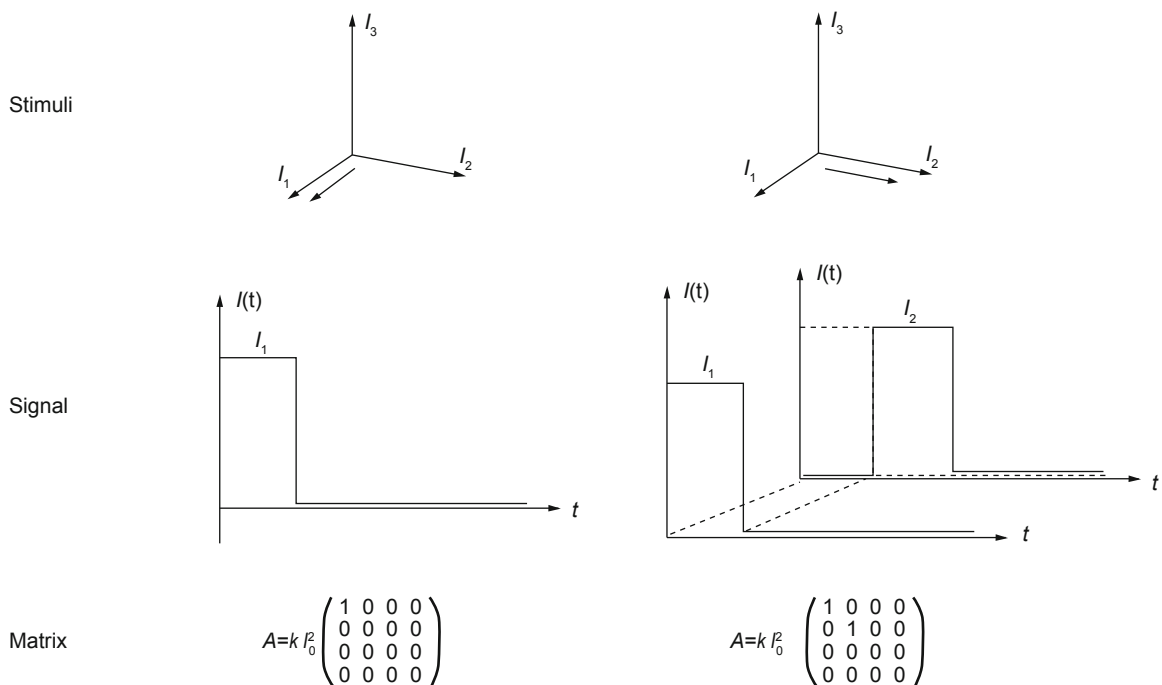

Fig. 3 Optimum design for an inverted 5-spot with one central producer. Here injection is limited by I_0. Stimuli represents the direction in the injection space (we only display 3 of 4 dimensions) where the external stimuli is applied. The second line, Signal, shows injection as a function of time. The last line shows the resulting Fisher information matrix for the overall scheme. Only the first two steps of the input design construction are shown.

separated by a certain gap to the following larger eigenvalue. Thus, following the same calculation given above, first order perturbation theory shows in this case that the best choice will be designing an input signal with injection rates within the constraints and maximum projection onto the subspace associated to the minimum eigenvalue. In this way the single-shot problem in the general case can also be solved using only diagonalization and linear programming.

5 Conclusions

This work studied the input design problem in the context of waterflood dynamic data. This problem is of interest to enhance the informative content of the dataset when an operational gap to modify injection rates is available. A numerical scheme and generic features of the optimal injection scheme have been derived. In particular it was shown that synchronized bang-bang inputs solve the problem when the constraints are minimum and maximum injection rates. In the general case, where a limited total injection is considered, piece-wise constant signals are found optimal. In this case we explicitly show how to solve the problem for a small future time window using only matrix-diagonalization and linear programming.

Acknowledgements

References

Albertoni A and Lake L W. Inferring interwell connectivity only from well-rate fluctuations in waterfloods. SPE Reservoir Evaluation & Engineering. 2003. 6(1): 6-16 (paper SPE 83381PA)

Bjorck A. Numerical Methods for Least Squares Problems. Society for Industrial and Applied Mathematics. 1996

Boyd S and Vandenberghe L. Convex Optimization. Cambridge University Press. 2004

Box G E P, Jenkins G M and Reinsel G C. Time Series Analysis: Forecasting and Control (4th ed.). Wiley. 2008

Elsgolts L. Differential Equations and the Calculus of Variations. Mir Publishers, Moscow, 1980

Fedorov V V. Model Oriented Design of Experiments. Springer. 1997 (GAMS) General Algebraic Modeling System, see: www.gams.com

Kaviani D, Personal Communication. 2012

Ljung L. System Identification: Theory for the User. Prentice Hall. 1999

Morelli E. Practical Input Optimization for Aircraft Parameter Estimation Experiments. Ph.D. Dissertation. G. Washington University, Joint Inst. For Advanc. of Flight Sc. NASA-CR-191462. 1990

Moreno G and Lake L W. On the uncertainty of interwell connectivity estimations from the capacitance-resistance model. Petroleum Science. 2014. 11(2): 265-271

Nguyen A P, Lasdon L S, Lake L W and Edgar T F. Capacitance resistive model application to optimize waterflood in a West Texas Field. SPE Annual Technical Conference and Exhibition, 30 October-2 November, 2011, Denver, Colorado, USA (paper SPE 146984-MS)

Reid D B. Optimal Inputs for System Identification. Ph.D. Dissertation. Stanford University. NASA-CR-128173. 1972

Sayarpour M, Kabir C S and Lake L W. Field applications of capacitance-resistance models in waterfloods. SPE Reservoir Evaluation & Engineering. 2009. 12(6): 853-864 (paper SPE114983PA)

Weber D. The Use of Capacitance-Resistance Models to Optimize Injection Allocation and Well Location in Water Floods. Ph.D. Dissertation. The University of Texas at Austin. 2009

Yousef A A, Gentil P, Jensen J L and Lake L W. A capacitance model to infer inter-well connectivity from production- and injection-rate fluctuations. SPE Reservoir Evaluation & Engineering. 2006. 9(6): 630-646 (paper SPE 95322PA)

Zarrop M. Optimal Experiment Design for Dynamic System Identification. Springer-Verlag. 1979