E31 GENERATION OF LOW-ORDER RESERVOIR MODELS USING POD, EMPIRICAL GRAMIAN AND SUBSPACE IDENTIFICATION

R. Markovinovic*, E.L. Geurtsen*, T. Heijn* and J.D. Jansen[†]

* *Delft University of Technology, Department of Applied Earth Sciences*

[†] *Shell International Exploration and Production, Smart Well Solutions*

Abstract

In this paper we present a number of data-driven approaches to obtain non-linear low-order models of heterogeneous reservoirs. Relying on the ‘proper orthogonal decomposition’ (POD) method of snapshots, the proposed approaches scale-up an existing high-order dynamic reservoir model to a model of lower order. Combined with so-called ‘balanced’ model order reduction techniques for linear systems the POD methods yield an intuitively motivated and systematic procedure for construction of low-order models for complex, high-dimensional nonlinear systems. The reduced-order nonlinear models are obtained in way that resembles standard model reduction methods for linear systems, such as balanced truncation, by preserving only those features of the dynamics that are most relevant to the control design, i.e. the most controllable and observable states. In one of the methods matrices are obtained that can be interpreted as generalized gramians for the nonlinear system. In addition we propose the use of subspace identification techniques to generate a linearized minimal balanced realization based on reservoir input-output data (flow rates and pressures), rather than Taylor-like linearization of the original full-order nonlinear model. The main advantage of this approach is the possibility to create linearizations of commercial reservoir simulation models without having access to the code. Although the reduced-order models resulting from our approach are nonlinear, the methods used in their construction are inherently linear; they produce a linear coordinate transformation by constructing and decomposing matrices obtained from ‘numerical experiments’ (simulation).

1. Introduction

Our objective is to assess the feasibility of system-theoretical concepts to reduce high-order multi-phase reservoir models, containing typically 10^3 - 10^6 variables, to models of much lower order, typically containing 10^1 - 10^3 variables. Our research forms part of a wider effort focusing on the use of measurement and control theory to optimize the operation of ‘smart wells’, i.e. wells equipped with downhole measurement and control devices. In particular we address the use of smart wells for the optimization of water flooding in heterogeneous reservoirs using optimal control theory [1]. Design of an optimal water flooding strategy involves multiple simulations of the flooding process, which is computationally intensive if performed with a high-order reservoir model, but may become much more feasible when using low-order models. In addition, we are working on the development of strategies to optimize the flooding process during the producing life of the reservoir. This involves the development of controllers, using ‘real-time’ measurements of pressures and flow rates in the wells in combination with a mathematical reservoir model. Such model-based controllers need to be of relatively low-order to be of practical value, and one possible route to construct them is to start from low-order reservoir models that contain only those features of the dynamics that are most relevant to the control design.

The material presented is not completely new from a system-theoretical point of view, and is based on classical theory. However, we believe that these concepts have not yet been applied to

petroleum engineering problems. An application of POD to a groundwater problem was recently reported in [2], but unlike in our paper, only linear models were considered.

2. Reservoir model

To illustrate ideas behind the proposed model reduction approaches, we simulated two-phase (oil-water) flow in a horizontal reservoir using an 8 x 8 finite difference model; see Fig. 1. The finite difference model can be represented in the following standard (nonlinear, continuous, time-invariant and affine-input) state space form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) \quad , \quad \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)) = \mathbf{C}\mathbf{x}(t), \quad (2.1)$$

where $\mathbf{x}$ is a 2*64-dimensional state vector consisting of the pressure and saturation values in all the grid blocks, $\mathbf{u}$ is a 2*8-dimensional vector of control inputs consisting of the flow rates at both the injectors and the producers, while the 2*2*8-dimensional output vector $\mathbf{y}$ consists of pressure and saturation measurements in the same wells. The 32 x 128 constant matrix $\mathbf{C}$ is therefore just a selector matrix selecting these 'output states'. For details of the relationship between the state space and the traditional reservoir engineering notations see e.g. [1].

Figure 1. Reservoir schematic.

Reservoir parameters:

m	number of grid cells in x and y direction	8	[-]
h	grid cell height	10	[m]
$\Delta x, \Delta y$	x, y -grid block width	25	[m]
S_w	initial water saturation	0.2	[m ³ /m ³]
p	initial oil pressure	40e6	[Pa]
ϕ	porosity	[0.06, 0.29]	[m ³ /m ³]
k	permeability	[1.7e-14, 6.7e-13]	[m ²]

Fluid parameters:

c	compressibility	1e-9	[1/Pa]
μ	viscosity	1e-3	[Pa.s]
ρ_o	density oil	8e2	[kg/m ³]
ρ_w	density water	1e3	[kg/m ³]

Table 1. Reservoir and fluid parameters.

3. Model reduction

3.1 Coordinate transformation

To reduce the order of a model generally means the finding a coordinate transformation of the state space that yields a meaningful ranking of the state components in terms of their influence on the system behavior. Those state components of the transformed system that have relatively little influence can then be truncated from the system description, without substantially degrading the system's predictive capabilities, i.e. its correctness. Mathematically this means the following. Consider a general nonlinear, continuous, time-invariant systems in state-space form,

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \quad , \quad \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)) \quad (3.1)$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^p$, and $\mathbf{y} \in \mathbb{R}^m$. Let $\Phi_k \in \mathbb{R}^{n \times k}$ be an transformation matrix with $k \ll n$ and $\text{rank}(\Phi_k) = k$, and write $\mathbf{x}(t) = \Phi_k \mathbf{z}(t) + \mathbf{r}(t)$. Substituting this in Eq. (3.1) yields

$$\Phi_k \dot{\mathbf{z}}(t) + \dot{\mathbf{r}}(t) = \mathbf{f}(\Phi_k \mathbf{z}(t) + \mathbf{r}(t), \mathbf{u}(t)) , \quad \mathbf{y}(t) = \mathbf{h}(\Phi_k \mathbf{z}(t) + \mathbf{r}(t), \mathbf{u}(t)) . \quad (3.2)$$

Using the concept of Galerkin projection one may require the residual $\mathbf{r}(t)$ to be orthogonal to the k -dimensional subspace on which the system dynamics is projected, or equivalently, to be orthogonal to the column space of Φ_k , i.e. $\Phi_k^T \mathbf{r}(t) = \mathbf{0}$ for all t . This results in

$$\Phi_k^T \Phi_k \dot{\mathbf{z}}(t) = \Phi_k^T \mathbf{f}(\Phi_k \mathbf{z}(t) + \mathbf{r}(t), \mathbf{u}(t)) , \quad \mathbf{y}(t) = \mathbf{h}(\Phi_k \mathbf{z}(t) + \mathbf{r}(t), \mathbf{u}(t)) , \quad (3.3)$$

or equivalently

$$\dot{\mathbf{z}}(t) = \Phi_k^\dagger \mathbf{f}(\Phi_k \mathbf{z}(t) + \mathbf{r}(t), \mathbf{u}(t)) , \quad \mathbf{y}(t) = \mathbf{h}(\Phi_k \mathbf{z}(t) + \mathbf{r}(t), \mathbf{u}(t)) , \quad (3.4)$$

where $\Phi_k^\dagger \triangleq (\Phi_k^T \Phi_k)^{-1} \Phi_k^T$ denotes the Moore-Penrose inverse of Φ_k , which exists since Φ_k has full column rank. Until now, the equation has remained exact. The k -dimensional approximation is based on the fact that the ‘projecting’ matrix Φ_k has been chosen such that the residual $\mathbf{r}(t)$ is minimized in some sense, so that the reduced, k^{th} -order model is given by

$$\dot{\mathbf{z}}(t) = \Phi_k^\dagger \mathbf{f}(\Phi_k \mathbf{z}(t), \mathbf{u}(t)) , \quad \mathbf{y}(t) = \mathbf{h}(\Phi_k \mathbf{z}(t), \mathbf{u}(t)) . \quad (3.5)$$

3.2 Proper Orthogonal Decomposition (POD)

It is obvious that the performance of a reduced order model depends on the choice of the basis vectors for a coordinate transformation of the state space. One approach to find a suitable basis for the desired coordinate transformation is application of the proper orthogonal decomposition (POD)¹ (see, e.g. [3-6]). A big advantage of the POD methodology is that it is physically motivated and the resulting empirical eigenfunctions give physical intuition to the behavior of the system. This in contrast to discretization methods for partial differential equations (PDEs), like finite difference, finite element or spectral methods, where basis functions are used that have very little connection to the problem at hand or to the underlying PDE.²

Consider a general nonlinear state-space model (3.1). In the POD method, the basis vectors are obtained from an ensemble of state time-series data (state snapshots), generated either by simulation or experiment, and taken at various time instants. For $N \gg n$, let $\mathbf{X} := [\mathbf{x}(1) \dots \mathbf{x}(N)]$ be the data matrix of state snapshots, and let $\mathbf{R} = \mathbf{X}\mathbf{X}^T$ be the corresponding covariance data matrix. It can be shown that the eigenvectors of the following eigenvalue problem:

$$\mathbf{R}\mathbf{v} = \lambda \mathbf{v} \quad (3.6)$$

determine principal directions in $\mathbb{R}^n$ that best fit the collected state snapshots, measured in terms of the relative ‘energy’, i.e. mean square fluctuations, associated with particular directions in the state space. The principal eigenvectors can be computed very reliably as the columns of the matrix Φ in the singular value decomposition

$$\mathbf{X} = \Phi \Sigma \Psi^T . \quad (3.7)$$

¹ Also known as ‘Karhunen-Loeve decomposition’, ‘principal component analysis’, ‘empirical orthogonal decomposition’ or ‘method of empirical eigenfunctions’.

² In finite element methods, for example, piecewise polynomials, in finite difference methods grid functions, and in spectral methods Legendre polynomials are often used.

3.3 Model reduction aspects

Let $\Phi_k = \Phi(:, 1:k)$, i.e. Φ_k consist of the first k columns of the orthogonal matrix Φ from the SVD (3.7). Following the general discussion from the beginning of this section, an ‘optimal’ reduced k^{th} -order model of (3.1) is given by³

$$\dot{\mathbf{z}}(t) = \Phi_k^T \mathbf{f}(\Phi_k \mathbf{z}(t), \mathbf{u}(t)), \quad \mathbf{y}(t) = \mathbf{h}(\Phi_k \mathbf{z}(t), \mathbf{u}(t)). \quad (3.8)$$

The optimality of the model is in the sense that a truncated series representation of the data using basis functions obtained by the POD method has a smaller mean square error than a representation by any other basis of the same dimension; see e.g. [3].

As the above review shows, this standard POD procedure concerns only the representation of the state snapshots, and the performance of the reduced-order model depends strongly on the choice of control inputs, boundary conditions (BCs) and initial conditions (ICs) used to generate the empirical time-series data. For nonlinear systems, small perturbations in these parameters can lead to qualitative different system responses. This is the main shortcoming of the method when applied on systems possessing the feature that principal components of the flow are (very) sensitive to the choice of control inputs, BCs or ICs. At first sight, this seems not to be the case for reservoir models, at least not for the ones we considered, where seemingly any set of empirically determined eigenfunctions will likely be efficient for model reduction purposes (see also the simulation results in Section 5). However, from the control-optimization point of view we propose consideration of alternative methodologies that take into account state-to-output behavior of the system as well. The standard POD method described above does not do this and its efficiency may therefore be diminished when constructing ‘black-box’ input-output models of the system. All of the alternative POD methods discussed below combine in one way or another the ideas of the standard POD and the so-called method of balanced state space realizations, to be explained next.

Remark. In addition to the shortcoming concerning the choice of control inputs, BCs and ICs mentioned above, the data collected while applying the standard POD may fail to capture dynamical effects occurring at widely different time scales. A representative example of such systems is obviously a multi-phase reservoir system, where the pressure part of the state-space has completely different time constants than the saturation part. To deal with this, one may wish to do sampling non-uniformly in time. The problem is, however, that usually not all of the time constants of the system are known a priori. A possibility to get a quick impression about them is to apply subspace identification, which is discussed in Section 4.

3.4 Balanced model order reduction

Consider a linear time invariant multi-input multi-output (MIMO) system of the form:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \quad (3.9)$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^p$ and $\mathbf{y} \in \mathbb{R}^m$ are the *state*-, the *input*-, and the *output* of the system, and $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are constant matrices of compatible dimensions. In the sequel we take the system (3.9) to be asymptotically stable and of minimal degree of freedom (i.e. both controllable and observable [7]). It can be shown (see e.g. [8]) that the minimum input energy required to bring the system state from $\mathbf{x}(-\infty) = \mathbf{0}$ to $\mathbf{x}(0) = \mathbf{x}_0$ is given by $\mathbf{x}_0^T \mathbf{W}_c^{-1} \mathbf{x}_0$, while the output energy

³ In this equation the fact is used that the orthogonality of the matrix Φ implies that $\Phi^\dagger = \Phi^T$.

produced by the (perturbed) state $\mathbf{x}_0$ equals $\mathbf{x}_0^T \mathbf{W}_o \mathbf{x}_0$, where the positive definite matrices $\mathbf{W}_c$ and $\mathbf{W}_o$ denote respectively the *controllability*- and *observability gramian* of the system, defined as

$$\mathbf{W}_c := \int_0^\infty e^{A^t} \mathbf{B} \mathbf{B}^T e^{A^T t} dt, \quad \mathbf{W}_o := \int_0^\infty e^{A^T t} \mathbf{C}^T \mathbf{C} e^{A t} dt. \quad (3.10)$$

Herein, $e^{A^t} \mathbf{B} \equiv \mathbf{x}_\delta(t)$ is the impulse state-response of (3.9), whereas $e^{A^T t} \mathbf{C}^T \equiv \mathbf{z}_\delta(t)$ is the impulse state-response solution of the dual system of (3.9), i.e. $\dot{\mathbf{z}}(t) = \mathbf{A}^T \mathbf{z}(t) + \mathbf{C}^T \mathbf{u}_d(t)$, $\mathbf{y}_d(t) = \mathbf{B}^T \mathbf{z}_d(t)$. Numerically $\mathbf{W}_c$ and $\mathbf{W}_o$ are found by solving the following Lyapunov equations:

$$\mathbf{A} \mathbf{W}_c + \mathbf{W}_c \mathbf{A}^T = -\mathbf{B} \mathbf{B}^T, \quad \mathbf{A}^T \mathbf{W}_o + \mathbf{W}_o \mathbf{A} = -\mathbf{C}^T \mathbf{C}. \quad (3.11)$$

It is a well-known result from system theory that a given input-output linear system is *similarity invariant* with respect to state space realizations like the one in (3.9). However, the input-state and the state-output relationships do differ from one realization to another. The main idea behind the balancing technique for model order reduction is to find a realization for (3.9), which corresponding controllability and observability gramians are equal and diagonal [8]: $\mathbf{W}_c = \mathbf{W}_o = \mathbf{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$, where $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ are the so-called *Hankel singular values* of (2.1) and are the square roots of the eigenvalues of the product $\mathbf{W}_c \mathbf{W}_o$. It is easily shown that these are similarity invariants. It is now clear that, in a balanced model, the state components in $\mathbf{x}_0$ corresponding to small Hankel singular values require much energy to be reached, while at the same time producing little energy on the output. Therefore, if there is a $s \in \{1, \dots, n-1\}$ for which $\sigma_s \gg \sigma_{s+1}$, the state components corresponding to $\sigma_m, m = s+1, \dots, n$ can be truncated from the system description. It should here also be emphasized that there are certain guaranteed error bounds in this approximation that are a priori known.

3.5 Empirical gramians balanced model reduction

As we have seen above, the central role in a balancing procedure for linear systems is played by the controllability and observability subspaces of the system. For nonlinear systems, however, while it is relatively straightforward to obtain an approximation to the controllability subspace using the POD eigenfunctions obtained from the simulation of the nonlinear system (3.1), difficulties arise with the approximation of the observability subspace. The reason is that the concept of a dual system does not really exist⁴ in a nonlinear setting. What one can try is to linearize the nonlinear system (3.1) and formulate the dual linearized system (the adjoint), from which an estimate of the observability gramian could be taken.

Another possibility is to follow the method developed in [10], where an ‘empirical’ observability gramian is defined, based on system outputs for various initial conditions. Below a short outline of the method is given. For a more comprehensive treatment of the method we refer to [10]. Consider again the general nonlinear system (3.1), and define the following sets:

$\mathcal{T} := \{\mathbf{T}_1, \dots, \mathbf{T}_r\}$ - a set of r orthogonal $n \times n$ matrices;

$\mathcal{M} := \{c_1, \dots, c_s\}$ - a set of s positive constants;

$\mathcal{E} := \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ - a set of n standard unit vectors in $\mathbb{R}^n$.

⁴ For extensions of the linear ‘adjoint’ characterization to nonlinear systems we refer to [9] and references therein.

For exponentially stable systems we have the following definitions of the ‘empirical’ controllability and observability gramian, respectively [10].

Definition 1. For the system (3.1), let $\mathcal{T}$, $\mathcal{E}$ and $\mathcal{M}$ be given sets as described above. The *empirical controllability gramian* $\hat{\mathbf{W}}_c$ is defined by

$$\hat{\mathbf{W}}_c = \sum_{l=1}^r \sum_{m=1}^s \sum_{i=1}^p \frac{1}{rsc_m^2} \int_0^\infty \Upsilon^{ilm}(t) dt, \quad (3.12)$$

where $\Upsilon^{ilm}(t) \in \mathbb{R}^{n \times n}$ is given by $\Upsilon^{ilm}(t) := (\mathbf{x}^{ilm}(t) - \mathbf{x}_{av}^{ilm})(\mathbf{x}^{ilm}(t) - \mathbf{x}_{av}^{ilm})^T$, and $\mathbf{x}^{ilm}(t)$ is the state of (3.1) corresponding to the impulse input $\mathbf{u}(t) = c_m \mathbf{T}_l \mathbf{e}_i \delta(t)$.

Definition 2. For the system (3.1), let $\mathcal{T}$, $\mathcal{E}$ and $\mathcal{M}$ be given sets as described above. The *empirical observability gramian* $\hat{\mathbf{W}}_o$ is defined by

$$\hat{\mathbf{W}}_o = \sum_{l=1}^r \sum_{m=1}^s \frac{1}{rsc_m^2} \int_0^\infty \mathbf{T}_l \Psi^{ilm}(t) \mathbf{T}_l^T dt, \quad (3.13)$$

where $\Psi^{ilm}(t) \in \mathbb{R}^{n \times n}$ has elements $\Psi_{ij}^{ilm}(t) := (\mathbf{y}^{ilm}(t) - \mathbf{y}_{av}^{ilm})(\mathbf{y}^{jlm}(t) - \mathbf{y}_{av}^{jlm})^T$, and $\mathbf{y}^{ilm}(t)$ is the output of (3.1) corresponding to the initial state $\mathbf{x}_0 = c_m \mathbf{T}_l \mathbf{e}_i \delta(t)$, with $\mathbf{u}(t) \equiv \mathbf{0}$ for $t \geq 0$.

It can be shown [10] that, when applied to linear systems (3.9), these empirical gramians are just the usual linear controllability and observability gramians. The gramians (3.12) and (3.13) may therefore be seen as computable generalizations of the linear gramians for nonlinear systems. Having obtained these empirical gramians, they need to be balanced, i.e. a non-singular transformation $\mathbf{T} \in \mathbb{R}^{n \times n}$ has to be found such that the transformed gramians $\mathbf{T} \hat{\mathbf{W}}_c \mathbf{T}^T$ and $\mathbf{T}^{-T} \hat{\mathbf{W}}_o \mathbf{T}^{-1}$ are equal and diagonal. A simple numerical technique to find such a $\mathbf{T}$ is as follows:

- Apply Cholesky factorization to $\hat{\mathbf{W}}_c$, i.e. $\hat{\mathbf{W}}_c = \mathbf{G} \mathbf{G}^T$, with $\mathbf{G} > \mathbf{0}$ lower triangular;
- Let $\mathbf{U} \Sigma^2 \mathbf{U}^T$ be a SVD of $\mathbf{G}^T \hat{\mathbf{W}}_o \mathbf{G}$. Then we have $\mathbf{G}^T \hat{\mathbf{W}}_o \mathbf{G} = (\mathbf{U} \Sigma^{1/2}) \Sigma (\mathbf{U} \Sigma^{1/2})^T$, or equivalently, $\Sigma = (\Sigma^{1/2} \mathbf{U}^T \mathbf{G}^{-1})^{-T} \hat{\mathbf{W}}_o (\Sigma^{1/2} \mathbf{U}^T \mathbf{G}^{-1})^{-1}$.
- Finally, take $\mathbf{T} := \Sigma^{1/2} \mathbf{U}^T \mathbf{G}^{-1}$ which yields $\mathbf{T} \hat{\mathbf{W}}_c \mathbf{T}^T = \mathbf{T}^{-T} \hat{\mathbf{W}}_o \mathbf{T}^{-1} = \Sigma$, as desired.

4. Subspace identification

For (very) large-scale systems, the above procedure, although more accurate, can be computationally too expensive to be of practical usage. In that case, one could try using the observability gramian matrix of a linear description of the model. However, the mathematical models used in many of the commercial reservoir simulators are often inaccessible to the user. Therefore we propose the use of so-called ‘subspace-based identification algorithms’ to generate linearized state-space reservoir descriptions, using only input-output (flow rates and pressures) data. Even if a simulator is accessible, so that one can linearize the system equations by some Taylor-linearization procedure, there could still be difficulties. For example solving of the Lyapunov equations (3.11) for very large systems can be computationally very demanding and moreover, numerically unreliable.

Having their origin in state-space realization theory, see [11], subspace-based system identification methods are in particular able to provide state-space models for multiple-input multiple-output (MIMO) linear systems directly from input-output data. Contrary to the classical prediction error methods, no non-linear parametric optimization problems have to be solved. Due to the non-iterative nature of the algorithms, stagnation in local minima or convergence problems are avoided. Moreover, by incorporating a balancing step the identified model can always be in balanced coordinates. Although considerably different in implementation, all subspace algorithms consist basically of two steps. Depending on the algorithm, the first step estimates directly from input-output data: a) the column space of the extended observability matrix (as in MOESP algorithms [12, 13]), from which the minimal dimension of the phase-space and the system matrices are easily extracted using well-known numerical linear algebra algorithms such as SVD and LQ decomposition, or b) (explicit or implicit) a (Kalman filter) state sequence by doing orthogonal or oblique projections [14]. The second step then recovers (or estimates) the remaining part of the system. An excellent overview of different subspace-based identification techniques is given in [15]. For a recent application of subspace techniques in reservoir identification see [16].

5. Application

The POD method of snapshots was applied to the system described in Section 2, where also the physical parameters and dimensions of the reservoir are given (Table 1). Assumed porosity and permeability distributions are depicted in Fig. 2a. The time step used was 5 days, and the total duration was about 6 years, which resulted in an ‘almost water-saturated’ reservoir, see Fig. 6. A number of simulations was performed, each time with different input profiles for the generation of state snapshots and for the verification of the reliability of the chosen basis functions (Fig. 2b). In Fig. 3 and Fig. 4 typical responses obtained for pressures and saturations in the (output) cells (1,1) and (4,8) are given, where the first number represents the x -coordinate, and the second number the y -coordinate in Fig. 1. As expected, when taking more basis functions, the approximation of the state variables (pressures and saturations) becomes better. It is apparent that for an almost perfect reconstruction of the pressure values, three eigenfunctions are in most of the cases more than sufficient⁵ (which is in good accordance with Fig. 5 depicting the eigenvalues of the covariance matrix of the snapshots in (3.6)), whereas for a good approximation of the saturation values 16 modes are needed. As the pressures and saturations possess completely different time scales, collecting snapshots non-uniformly in time (which was not done in these simulations) would probably improve the reduction order even further. The uniformly collected snapshots may probably also be blamed for the fact that the reduced-order models obtained from balancing empirical gramians did not lead to a considerable improvement. Another cause may be the simplicity of the system considered. It is noted that the added value of the empirical gramians may become more prominent when reduced models are used for the design of control or optimization algorithms.

6. Concluding remarks

- Proper Orthogonal Decomposition (POD) is a promising technique to obtain low-order multi-phase reservoir models. The low-order models are derived from state vector ‘snapshots’ (pressures and saturations) generated with the original high-order model. POD-based reduced reservoir models are nonlinear, even although the techniques in deriving them are inherently linear.

⁵ Actually more than 99% of the energy contained in the snapshots is captured by just the first eigenvalue of the data covariance matrix.

- Improved reduction methods can be obtained through the combination of POD with balanced model order reduction techniques. In particular, we propose to combine POD with an empirical observability gramian based on system outputs for various initial conditions applied to the original nonlinear system. Alternatively one can use the observability gramian of the dual of a linearized form of the original nonlinear high-order model.
- Subspace identification can be used to obtain linearized system models from input-output data only. This allows for linearization of reservoir models in commercial simulators without access to the code.
- Application of the POD-based model reduction to a two-phase (oil-water) reservoir model produced promising preliminary results. The original model contained 128 state variables (64 pressures, 64 saturations). A satisfactory pressure response was obtained with 3 modes only, while an acceptable saturation response required 16 modes.

Figure 2. a) Permeability (left) and porosity (right) distributions in the reservoir; b) Input signal profile for generating snapshots (left) and that for verifying the reliability of reduced order models (right).

Figure 3. Pressure and saturation responses in cell (1,1) for the full-order nonlinear model (blue '—') and the reduced order-model (red '-') for respectively 3, 6, and 16 modes (from left to right). In each block of four sub-figures the left figures depict the pressure response, while the right figures depict the saturation response. Furthermore, the upper figures depict the response to the original simulation input signal, while the lower figures depict the response to the verification input.

Figure 4. Pressure and saturation responses in cell (4,8) for the full-order non-linear model (blue ‘—’) and the reduced order model (red ‘-’) for respectively 3, 6, and 16 modes. See Fig. 3 for extended key.

Figure 5. Eigenvalues of the data covariance matrix in (3.6).

Figure 6. Almost ‘water-saturated’ reservoir.

References

- [1] N. Dolle, D.R. Brouwer and J.D. Jansen. *Dynamic optimization of water flooding with multiple injectors and producers using optimal control theory*, Proc. XIVth Int. Conf. on Comput. Meth. in Water Resources (CMWR), Delft, The Netherlands, 2, 1435-1442 (2002).
- [2] P.T.M. Vermeulen, A.W. Heemink and C.B.M. Testroet. *Reduction of large-scale numerical ground water flow models*, Proc. XIVth Int. Conf. on Comput. Meth. in Water Resources (CMWR), Delft, The Netherlands, 1, 397-404 (2002).
- [3] G. Berkooz, P. Holmes, and J.L. Lumley. *The proper orthogonal decomposition in the analysis of turbulent flows*, Annual Review of Fluid Mechanics, 25, 539 – 575 (1993).
- [4] L. Sirovich, M. Maxey, and H. Tarman. *Eigenfunction analysis of turbulence correlations in thermal convection*, Bull. Amer. Phys. Soc., 31, 1676 (1986).
- [5] L. Sirovich. *Turbulence and the dynamics of coherent structures, Part I: Coherent structures*, Quart. Appl. Math., 45, 561-571 (1987).
- [6] L. Sirovich. *Turbulence and the dynamics of coherent structures, Part II: Symmetries and transformations*, Quart. Appl. Math., 45, 573-582 (1987).
- [7] R.E. Kalman. *Mathematical description of linear dynamical systems*, J. Control (SIAM), 1, 152-192 (1963).
- [8] B.C. Moore. *Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction*, IEEE Trans. Automat. Control, AC-26, No.1, 17-32 (1981).

- [9] K. Fujimoto, J.M.A. Scherpen and W.S. Gray. *Hamiltonian realizations of nonlinear adjoint operators*. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, Princeton, USA, (March, 2000).
- [10] S. Lall, J.E. Marsden and S. Glavaski. *A subspace approach to balanced truncation for model reduction of nonlinear control systems*, Int. J. Robust Nonlinear Control, **12**, 514-535 (2002).
- [11] B. Ho and R. Kalman. *Efficient construction of linear state variable models from input/output functions*, Regelungstechnik, **14**, 545-548 (1966).
- [12] M. Verhaegen and P. Dewilde. *Subspace model identification, part I: The output-error state-space model identification class of algorithms*, Int. J. Control, **56** (5), 1187-1210 (1992).
- [13] M. Verhaegen. *Identification of the Deterministic Part of MIMO State Space Models given in Innovations Form from Input-Output Data*, Automatica, **30** (1), 61-74 (1994).
- [14] P. van Overschee and B. De Moor. *A Unifying Theorem for Three Subspace System Identification Algorithms*, Automatica, **31** (12), 1853-1864 (1995).
- [15] M. Viberg. *Subspace-based Methods for the Identification of Linear Time-Invariant Systems*, Automatica, **31** (12), 1835-1851 (1995).
- [16] R. Markovinovic, E.L. Geurtsen and J.D. Jansen. *Subspace identification of low-order reservoir models*, Proc. XIVth Int. Conf. on Comput. Meth. in Water Resources (CMWR), Delft, The Netherlands, **1**, 281-288 (2002).