

Non-intrusive reduced-order modeling for multiphase porous media flows using Smolyak sparse grids

Dunhui Xiao^{1,2}, Zhi Lin³, Fangxin Fang^{1,*}, Christopher C. Pain¹, Ionel M. Navon⁴, Pablo Salinas¹ and Ann Muggeridge⁵

¹*Applied Modelling and Computation Group, Department of Earth Science and Engineering, Imperial College London, Prince Consort Road, London, UK*

²*China University of Geosciences, Wuhan 430074, China*

³*Zhejiang University, Hangzhou, China*

⁴*Department of Scientific Computing, Florida State University, Tallahassee, FL 32306-4120, USA*

⁵*Department of Earth Science and Engineering, Imperial College London, London, UK*

SUMMARY

In this article, we describe a non-intrusive reduction method for porous media multiphase flows using Smolyak sparse grids. This is the first attempt at applying such a non-intrusive reduced-order modelling (NIROM) based on Smolyak sparse grids to porous media multiphase flows. The advantage of this NIROM for porous media multiphase flows resides in that its non-intrusiveness, which means it does not require modifications to the source code of full model. Another novelty is that it uses Smolyak sparse grids to construct a set of hypersurfaces representing the reduced-porous media multiphase problem. This NIROM is implemented under the framework of an unstructured mesh control volume finite element multiphase model. Numerical examples show that the NIROM accuracy relative to the high-fidelity model is maintained, whilst the computational cost is reduced by several orders of magnitude. Copyright © 2016 John Wiley & Sons, Ltd.

Received 8 December 2015; Revised 29 April 2016; Accepted 13 May 2016

KEY WORDS: non-intrusive; Smolyak sparse grid; POD; multiphase; porous media

1. INTRODUCTION

The numerical simulation of porous media multiphase flows is very important, especially in petroleum reservoir simulation. However, in the context of uncertainty studies, sensitivity analyses or optimal design, hundreds or thousands of runs of the reservoir model are needed in order to analyse the parameters statistically. This high-computational cost renders the analysis almost impractical. In these cases, model reduction technology is a viable way to mitigate the computational cost as it offers the potential to simulate systems with substantially increased computation efficiency while maintaining accuracy.

A variety of model reduction methods have been proposed to ease the intensive computational cost in porous media multiphase flow problems. For example, reduced-physics models that simplify the physics [1, 2], upscaling methods that solve the system using a coarser grid [3], multiscale methods that solve the equations on coarse grids using basis functions captured on the fine grids [4, 5]

*Correspondence to: Fangxin Fang, Applied Modelling and Computation Group, Department of Earth Science and Engineering, Imperial College London.

†E-mail: f.fang@imperial.ac.uk

and snapshot-based methods that record results from solutions at each time step. The snapshot-based model reduction method are based on proper orthogonal decomposition (POD) and its variants [6].

Streamline methods are one type of reduced-physics model. The idea of the streamline method is that it decouples the flow and transport equations into a set of one dimensional problems and then solves it sequentially. It is used in reservoir simulations [7, 8] and history matching [9]. These reduced-physics models result in a considerable speed-up for some problems, however, is still necessary to solve for the pressure in the full model [6]. Upscaling methods allow the problem to be simulated on a much coarser mesh [10]. This can result in a considerable improvement in CPU time, but the methods themselves tend to be empirical and the resulting coarse grid model may not always reproduce the results obtained from a fine grid simulation [11, 12]. Multiscale models use different scale grids (fine and coarse) to discretise the underlying governing equations. This method has been reported to provide a factor of 10–20 speed-up in the finite-volume framework for reservoir simulation [13]. Reduced-basis methods are also used to approximate the coupled Stokes equations and darcy flow in porous media [14].

Reduced-order models are used widely in many engineering fields. POD is the most popular method in the framework of reduced-order modelling. It has proven to be an efficient means of deriving a reduced basis for high-dimensional problems and has been used in a variety of research areas including ocean modelling [15–22], convective Boussinesq flows [23], inversion problems [24], fluid mechanics [16, 17, 25], convection diffusion reaction equations [26], elastic plastic structural problem [27], shape optimisation [28], molecular dynamics simulation [29], fluid-structure interactions problems [30], porous media [31, 32] and air pollution [33]. It is also used combined with the trajectory piecewise linearisation method to simulate two-phase subsurface flows [34] and production optimisation [35].

In most cases, the source code that describes the physical system has to be modified in order to adapt to new requirements. These modifications can be complex, especially in legacy codes, or may be impossible (e.g. in commercial software) [36]. To circumvent this shortcoming, more recently, non-intrusive methods have been introduced into ROMs. Walton *et al.* [17] presented a method based on POD and radial basis function (RBF) interpolation and applied it to modelling unsteady fluid flows [37]. Xiao *et al.* also presented a non-intrusive reduced-order modelling (NIROM) based on POD-RBF [38]. Klie proposed a NIROM based on POD, RBF and the discrete empirical interpolation method [39]. However, most NIROMs still suffer from computational costs because of the exponential increase of the number of discretisation variables. To cope with this 'curse of dimensionality', the Smolyak sparse grid interpolation method [40] was introduced to NIROM [41]. In this paper, we apply Smolyak sparse grid interpolation to the non-intrusive reduced-order modelling of porous media flows for the first time. The NIROM is implemented under the framework of a high-fidelity unstructured finite element mesh fluid model (Multiphase-Fluidity, developed by Applied Modelling and Computation Group at Imperial College London) [42]. In this NIROM, a number of snapshots are recorded at regular time intervals from numerical solutions. The POD method is then used to generate a small number of basis functions. Using the Smolyak sparse grid method, a set of interpolation functions (hypersurface) is constructed to represent the reduced system. Having obtained the hypersurface, solutions of ROM for current time step can be calculated by inputting results from previous time steps.

The structure of the paper is as follows. Section 2 describes the governing equations. Section 3 presents the standard POD method. Section 4 derives the methods of constructing a NIROM using the POD and Smolyak sparse grid method for multiphase porous media problems. Section 5 demonstrates the method's capabilities in three porous media multiphase flow test cases. Finally, in section 6, the summary and conclusions are drawn.

2. GOVERNING EQUATIONS

The unstructured finite-element mesh model (Multiphase-Fluidity) is applied to the Darcy's law for immiscible porous media multiphase flows. Assume that a volume of an incompressible porous medium contains N_p immiscible fluid phases, and each phase V_η assumes the form of,

$$V_\eta = \psi S_\eta, \quad (1)$$

where the subscript, η , denotes the phase, and S_η is the saturation of that phase, ψ is porosity of porous medium and $\sum_{\eta=1}^{N_p} S_\eta = 1$. Darcy's Law has the following form [43],

$$\mathbf{q}_\eta = -\frac{k_{r\eta}}{\mu_\eta} \mathbf{K} (\nabla P_\eta - \rho_\eta \mathbf{g}), \quad (2)$$

where \mathbf{q}_η denotes a volumetric flux rate (Darcy velocity) of phase η ; $k_{r\eta}$ denotes the relative permeability of the phase η and is a function of S_η ; μ_η denotes the dynamic viscosity of the fluid phase η ; P_η denotes the fluid pressure of phase η ; \mathbf{K} denotes a second rank tensor describing the permeability of the porous medium; ρ_η denotes the mass density of phase η ; and \mathbf{g} denotes the gravitational acceleration vector.

The Darcy velocity can be rewritten as in terms of the (average or interstitial) fluid velocity \mathbf{v}_η ,

$$\mathbf{q}_\eta = V_\eta \mathbf{v}_\eta = \psi S_\eta \mathbf{v}_\eta. \quad (3)$$

Then, Equation 2 can be rewritten as,

$$S_\eta \mathbf{\Lambda}_\eta \mathbf{v}_{\psi\eta} = -\nabla P_\eta + \rho_\eta \mathbf{g}, \quad (4)$$

where

$$\mathbf{v}_{\psi\eta} = \psi \mathbf{v}_\eta, \quad (5)$$

and

$$\mathbf{\Lambda}_\eta = \frac{\mu_\eta}{k_{r\eta}} \mathbf{K}^{-1}. \quad (6)$$

3. PROPER ORTHOGONAL DECOMPOSITION

For the POD method, a new set of modes (basis functions) is constructed from a collection of snapshots E recording the solutions of high-fidelity full model at a regular time intervals. In the POD formulation presented here, snapshots of each solution field, that is, velocity components solutions \mathbf{u}_x , \mathbf{u}_y , \mathbf{u}_z , pressure solutions p and saturation solutions S are recorded individually. The fields are stored in separate snapshot matrices $E_{u,x}$, $E_{u,y}$, $E_{u,z}$, E_p and E_S . Because each snapshot matrix is treated in an identical way, a generic snapshot matrix E is used here to describe the POD formulation for the sake of simplicity.

Taking the mean of snapshots from the snapshots matrix,

$$\tilde{E} = E - \bar{E}, \quad (7)$$

where

$$\bar{E} = \frac{1}{N_s} \sum_{j=1}^{N_s} E_{N_s}. \quad (8)$$

A set of basis functions $\{\phi\}$ is obtained by means of the POD method. This involves performing a singular value decomposition of the snapshot matrix \tilde{E} given in the form,

$$\tilde{E} = U \Sigma V^T. \quad (9)$$

The terms U and V denote the matrices that consist of the orthogonal vectors for $\tilde{E} \tilde{E}^T$ and $\tilde{E}^T \tilde{E}$, respectively, (Σ being a diagonal matrix). The singular values of snapshot matrix \tilde{E} are listed in order of decreasing magnitude in matrix Σ . A set of POD basis functions can be obtained by the first m columns of matrix U ,

$$\Phi_i = U_{:,i}, i \in \{1, 2 \dots m\} \quad (10)$$

The POD basis functions are optimal in the sense that these vectors can be the closest to the snapshot matrix \tilde{E} in Frobenius norm.

In POD formulation, any variable ψ (for example, the velocity, pressure and saturation components) can be expressed by,

$$\psi = \bar{\psi} + \sum_{j=1}^P \alpha_j \phi_j, \quad (11)$$

where α_j denotes the j^{th} POD coefficient, and $\bar{\psi}$ is the mean of the snapshots for the variable ψ .

4. THE SMOLYAK SPARSE GRID METHOD FOR CALCULATING THE PROPER ORTHOGONAL DECOMPOSITION COEFFICIENTS

In this section, the process of calculating the solutions of reduced-order multiphase porous media model (POD coefficients) is described. The Smolyak sparse grid interpolation method is used to construct a set of hypersurfaces representing the reduced-multiphase porous media system. Then, we use those hypersurfaces to calculate the solutions of the ROM.

The Smolyak sparse grid interpolation is an efficient method to cope with the high-dimensional problems [40]. It has been used in a number of research fields such as stochastic collocation approach for solving partial differential equations [44], model reduction problems [45]. In the work of [45], sparse grid is combined with a multilevel greedy method to derive a reduced-basis method in order to deal with the intensive computational cost.

An approximation of a d -dimensional function in a full tensor product form can be written as

$$\left(U^{l_1} \otimes \dots \otimes U^{l_d} \right) (f) = \sum_{i_1=1}^{O_{l_1}} \dots \sum_{i_d=1}^{O_{l_d}} f \left(x_{i_1}^{l_1}, \dots, x_{i_d}^{l_d} \right) \cdot \left(H_{i_1}^{l_1} \otimes \dots \otimes H_{i_d}^{l_d} \right), \quad (12)$$

where $O_{l_1}, O_{l_2} \dots O_{l_d}$ denote the number of knots in each dimension; $f \left(x_{i_1}^{l_1}, \dots, x_{i_d}^{l_d} \right)$ represent the function values at $\left(x_{i_1}^{l_1}, \dots, x_{i_d}^{l_d} \right)$; $O_{l_d} = 2^{l_d-1} + 1$; O_l ($l \in \mathbb{N}$) denotes the number of knots; $H_i^l \in C([-1, 1])$ is the approximation formulas and $f \left(x_i^l \right)$ denotes the value of the function f at x_i^l .

The disadvantage of the full tensor product approximation earlier is that it requires $O_{l_1} \times \dots \times O_{l_d}$ function values on the full tensor product grid. The number of function values increases exponentially with dimensions size d . In this case, the Smolyak sparse grid approach is used to cope with the 'curse of dimensionality' problem and has the form of

$$A(q, d) = \sum_{q-d+1 \leq |\mathbb{I}| \leq q} (-1)^{q-|\mathbb{I}|} \cdot \binom{d-1}{q-|\mathbb{I}|} \left(U^{l_1} \otimes \dots \otimes U^{l_d} \right), \quad (13)$$

where $|\mathbb{I}| = l_1 + \dots + l_d$, $\left(U^l \right) (f) = \sum_{i=1}^{O_l} f \left(x_i^l \right) \cdot \left(H_i^l(x) \right)$, the spatial dimensional size d equals to the size of the reduced-order space (i.e. the number of POD bases m). A set of Smolyak sparse grid interpolation functions $A_j(q, d)$ is then used to estimate the POD coefficient α_j^n at time level $n + 1$,

$$\alpha_j^{n+1} = \hat{f}_j \left(\alpha_1^n, \alpha_2^n, \dots, \alpha_m^n \right), \quad j \in \{1, 2, \dots, m\}, \quad (14)$$

where \hat{f} is the hypersurface calculated by equation (13).

The algorithmic process of constructing a set of hypersurfaces and calculating the solutions of ROM can be posed as follows:

- (a) Choose a number of sparse interpolation nodes $\alpha^{r,0} = (\alpha_1^{r,0}, \alpha_2^{r,0}, \dots, \alpha_m^{r,0})$ (where $r \in \{1, 2, \dots, R\}$ on the full tensor product grid, R is the number of sparse points to be chosen and is much less than the number of nodes of the full tensor product grid, 0 denotes the initial time step), which lie in the product interval

$$[A_{min}, A_{max}] = [\alpha_{1,min}, \alpha_{1,max}] \cdots \otimes [\alpha_{j,min}, \alpha_{j,max}] \cdots \otimes [\alpha_{m,min}, \alpha_{m,max}]$$

(where $\alpha_{j,min}$ and $\alpha_{j,max}$ are the minimum and maximum values of the j^{th} POD coefficient, $j \in \{1, 2, \dots, m\}$);

- (b) Calculate a corresponding set of the function values $\alpha_j^{r,1} = f_j(\alpha^{r,0})$ located at the sparse nodes $\alpha^{r,0}$ through running the full model one time step from time level 0 to time level 1:
(c) After obtaining a set of $\alpha_j^{r,1}$, then construct a set of interpolation functions $A_j(q, d)$, $j \in \{1, 2, \dots, m\}$ using Equation 13;
(d) Calculate the current time step's POD coefficient α^{n+1} using the interpolation formula (14) by inputting POD coefficient α^n at previous time step.
(e) Calculate the velocity, pressure and saturation solutions \mathbf{u}^n , p^n and S^n on full space at current time step $n + 1$ by projecting $\alpha_{\mathbf{u},j}^n$, $\alpha_{p,j}^n$ and $\alpha_{S,j}^n$ onto the full space.

$$\mathbf{u}^{n+1} = \sum_{j=1}^m \alpha_{\mathbf{u},j}^n \Phi_{\mathbf{u},j}, \quad p^n = \sum_{j=1}^m \alpha_{p,j}^n \Phi_{p,j}, \quad S^n = \sum_{j=1}^m \alpha_{S,j}^n \Phi_{S,j},$$

The advantage of using the Smolyak sparse grid interpolation method is that the interpolation function values need to be determined only at the sparse grid mesh points rather than at all the points on the full tensor product grid, thus resulting in an impressive computational economy.

5. NUMERICAL EXAMPLES

A demonstration of application of the NIROM based on POD and Smolyak sparse grid method to porous media multiphase is presented in this section. Three porous media multiphase test cases are 1D Buckley–Leverett, 2D Buckley–Leverett and a fracture problems. The test cases were simulated at the platform of the framework of an unstructured mesh control volume finite element (CVFEM) multiphase model, Multiphase-Fluidity [42]. The solutions from the fidelity full model provided the exact solutions for model comparison, as well as the snapshots for the POD basis function generation. An error analysis was carried out comprising the root mean square error (RMSE) and correlation coefficient, which considers all the nodes on the full finite element mesh and all the time steps.

5.1. Case 1: 1D Buckley–Leverett test case

The first application is a water-flooding test case. Water-flooding is a widely technique in oil reservoir engineering. As illustrated in Figure 1, the water is injected into the reservoir to increase the reservoir pressure, the oil is then displaced toward the production well [46]. The analytical solution (the Buckley Leverett solution) for a water-flooding in a 1D homogeneous reservoir predicts the formation of a self sharpening shock front between the water and the oil. This is challenging to model with a fine grid model let alone a ROM. In this case, the results of the NIROM are compared with the high-fidelity full model. From the high-fidelity full model simulation, with a mesh of 225 nodes, 120 snapshots were obtained by running the full order model and outputting all solution variables at regularly spaced time intervals of $\Delta t = 30$. The simulation period was dimensionless [0–3600], and a time step of $\Delta t = 3$ was used for all models.

Figure 2 shows the saturation solutions at time instances 1200 and 3600. The solutions compare the predictions from NIROM based on POD and Smolyak sparse grid using 12, 24 and 36 POD bases functions with full model.

Figure 3 compares the saturations predicted by the full model and the Smolyak sparse grid ROM at a particular position using 12, 24 and 36 POD bases. In order to demonstrate clearly the effects,

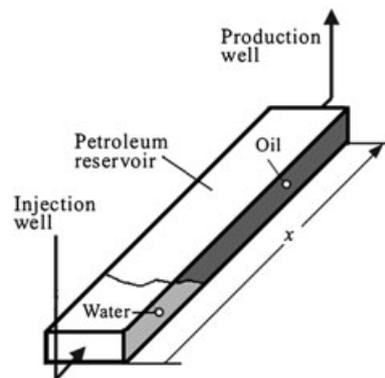


Figure 1. Water-flooding technique for oil production.

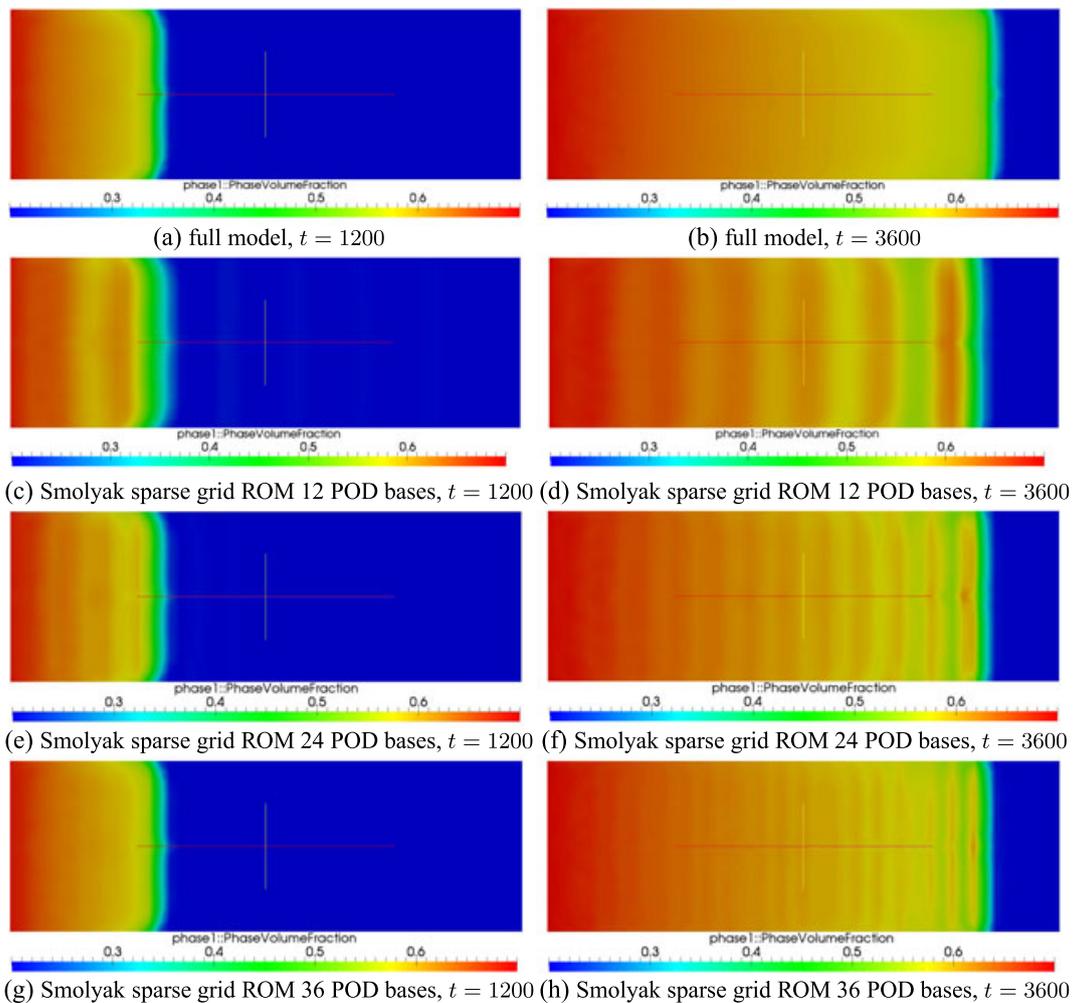


Figure 2. The figures displayed earlier show the saturation solutions of the Buckley–Leverett problem at time instances 1200 and 3600. The solutions compare the predictions from NIROM using 12, 24 and 36 POD bases functions with high-fidelity model.

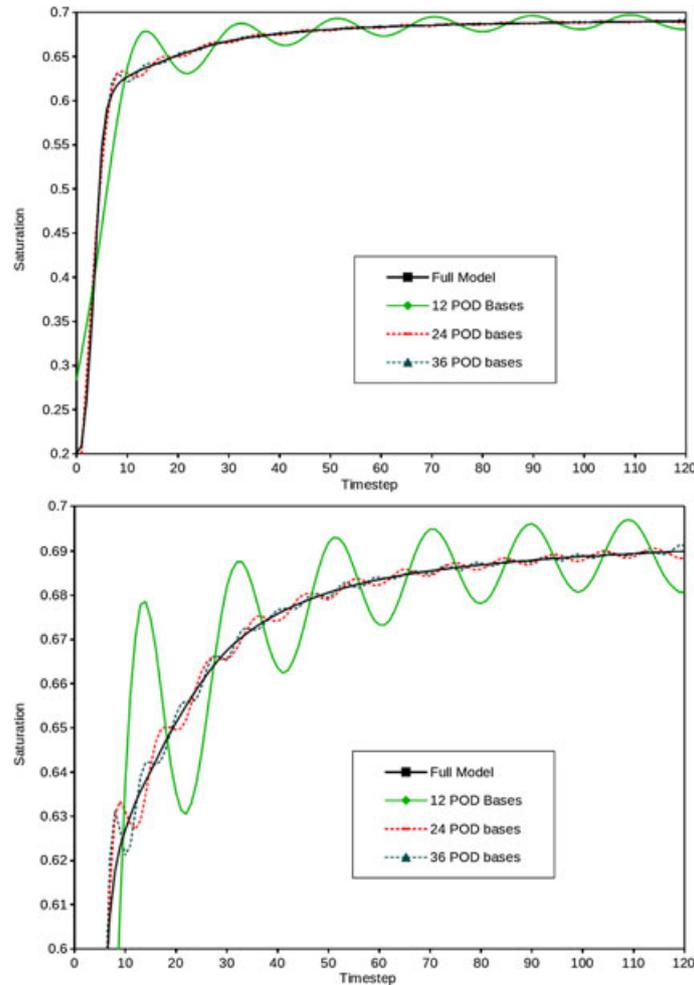


Figure 3. The graphs show the solution saturations predicted by the full model and the Smolyak sparse grid NIROM at a point($x=0.021$, $y=0.167$) using 12, 24 and 36 POD bases. The figure below the is a zoom of the figure above.

an enlarged figure is also provided, which shows the saturation values only between 0.6 and 0.7. As shown in the figure, there are some wiggles in the solutions. These wiggles are caused by the error of truncation of the basis functions. The wiggles are almost eliminated by using a larger number of POD basis functions.

Figure 4 shows some of the POD basis functions of this test case. They are the first, second, third, fourth, fifth, 12th, 24th and 36th POD bases functions, respectively. As shown in the figure, the 1st–4th POD basis functions capture most of the energy in the flows as well as major flow structures, the 1st–4th POD basis functions capture the details of small-scale flows.

Figure 5 shows the saturation profile at the first time step, the 50th time step, the 80th time step and the 120th time step. As can be seen from the figure, the results are closer to the high-fidelity full model when a larger number of POD basis functions is chosen.

5.2. Case 2: 2D High-permeability domain embedded in a low-permeability domain

The second case is comprised of a high-permeability domain embedded in a lower permeability domain. Water is injected over the left-hand boundary and oil and water are produced at a constant rate from the right hand face. From the full model simulation, with a mesh of 12 390 nodes, 50 snapshots were obtained by running full order model at regularly spaced time intervals of $\Delta t = 0.001$ for each solution variables. The simulation period is $[0 - 0.05]$, and a time step of $\Delta t = 0.0001$ is used for all models.

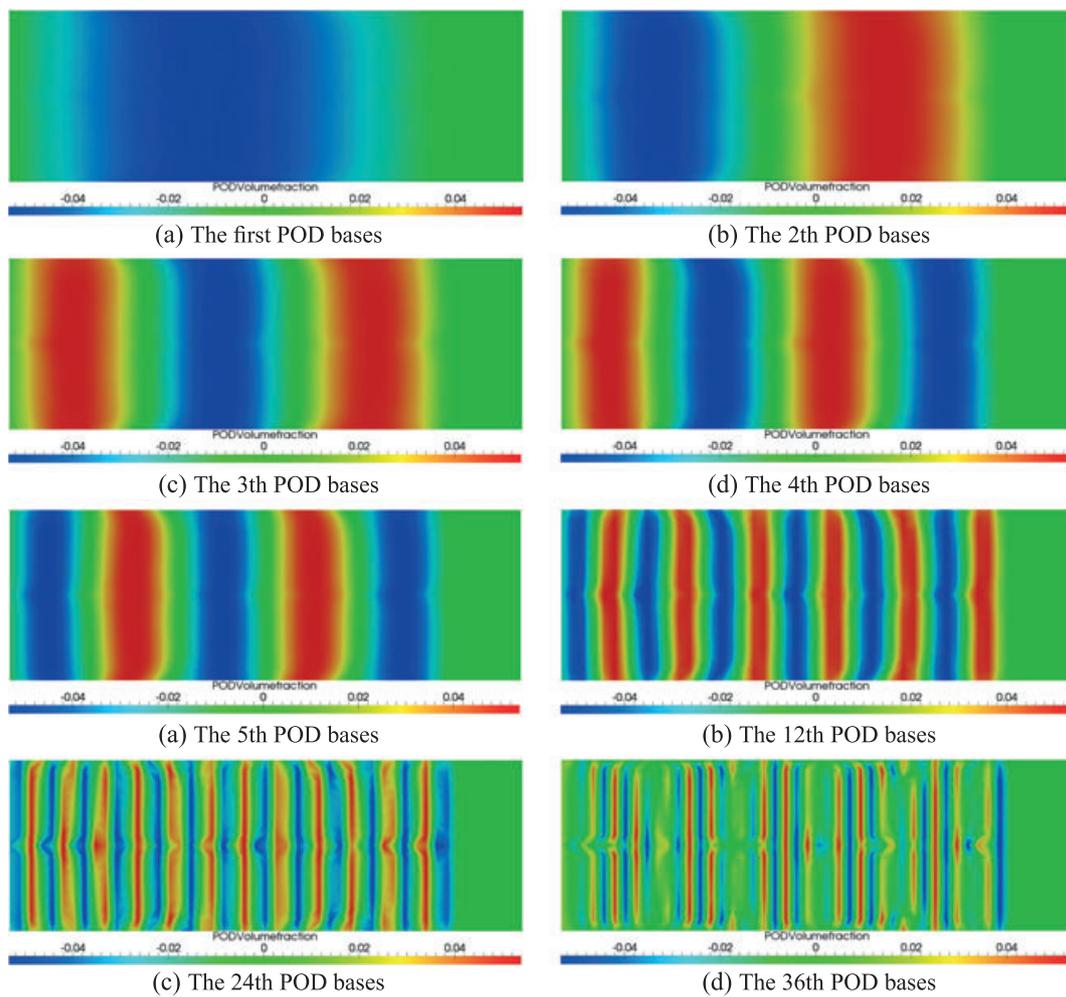


Figure 4. The figures displayed above show some of the POD bases functions of the 1-D Buckley-Leverett problem.

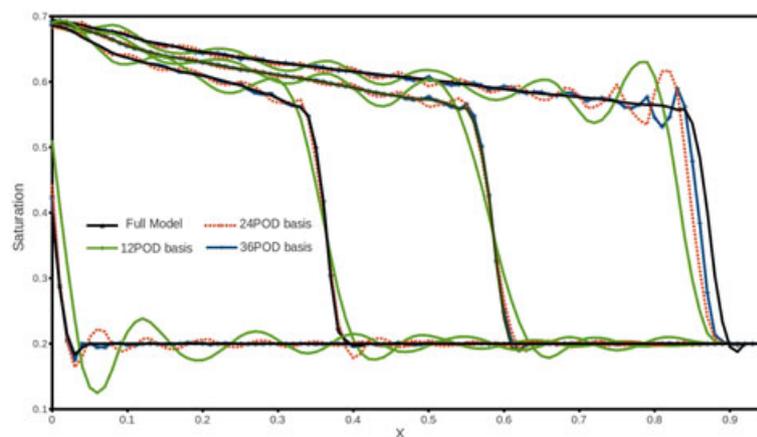


Figure 5. The graph shows the saturation profile from the full model and the NIROM using Smolyak sparse grid using 12, 24 and 36 POD bases at the first, 50th, 80th and 120th time interval.

Figure 6 shows the solutions of the 2D Buckley–Leverett problem at times 0.03 and 0.045. We compare the solutions from the NIROM based on POD and the Smolyak sparse grid using 12 and 36 POD basis functions with those from the full model. As we can see from the Figure 6(d), the water saturation in the high-permeability domain is not well predicted when 12 POD bases functions are used. This is improved by increasing the number of POD basis functions to 36, as shown in the

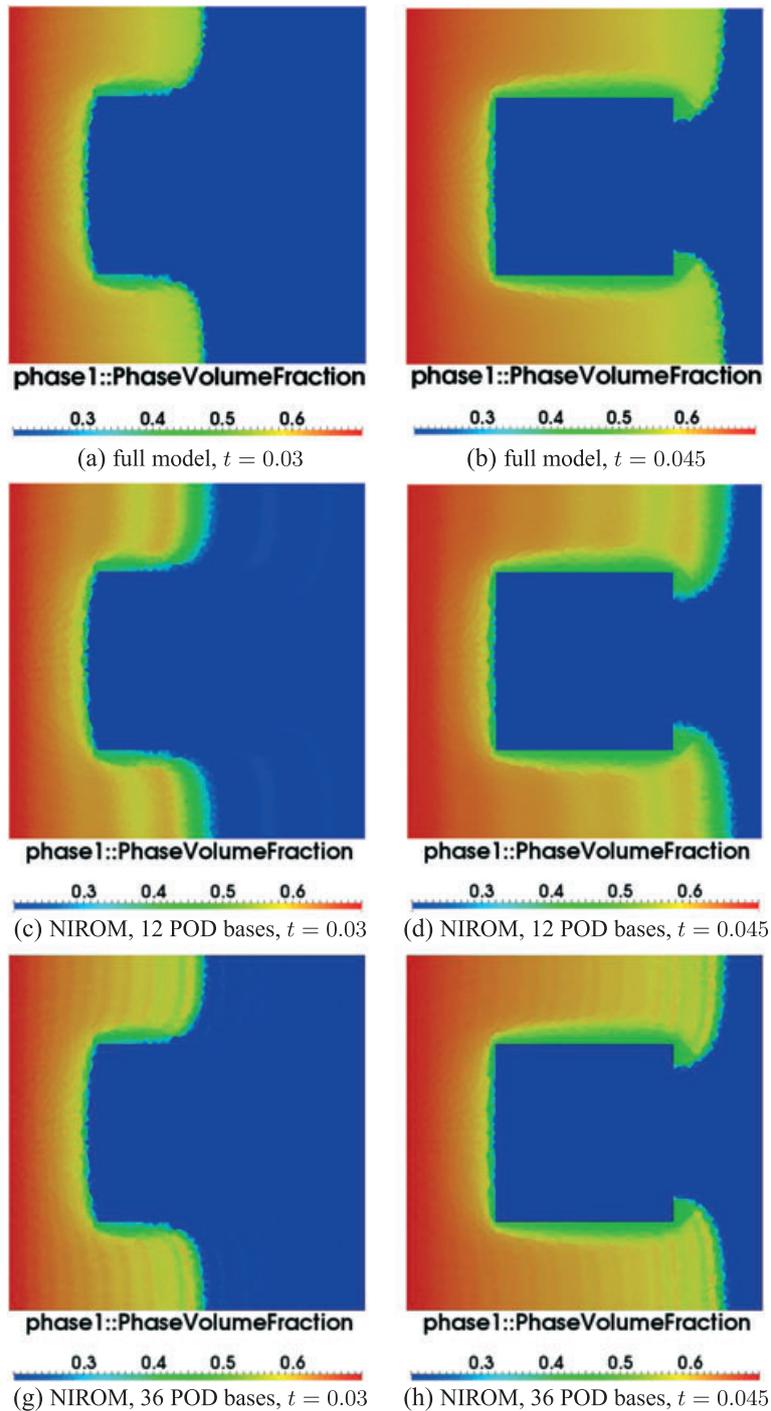


Figure 6. The figures displayed earlier show the solutions of the 2D Buckley–Leverett problem at time instances 0.03 and 0.045. The solutions compare the predictions from NIROM based on POD and Smolyak sparse grid using 12 and 36 POD bases functions, respectively, with the high-fidelity model.

Figure 6(f). To further demonstrate the capability of NIROMs, the model solution at a particular point along the domain is given in Figure 7. This figure shows the saturation solution predicted by the full model and the NIROM using 12 and 36 POD bases at the position ($x = 0.608, y = 0.924$) within the reservoir domain. The figure clearly illustrates that the solutions of NIROM using 36 POD basis functions are much more closer to those from the full model. They perform better than NIROM using 12 POD basis functions.

Table I shows a comparison of CPU (unit: s) required for running the full model and the NIROM for each time step. It can be seen that the NIROM is CPU time efficient once it has been constructed,

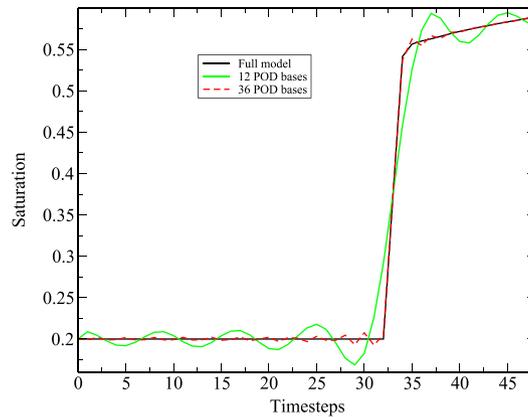


Figure 7. The graph shows the saturation solutions predicted by the full model, and the NIROM using 12 and 36 POD bases at a location ($x = 0.608, y = 0.924$).

Table I. Comparison of CPU (unit: s) required for running the full model and NIROM for each time step.

Model	Assembling and solving	Projection	Interpolation	Total
Full model	11.050	0.000	0.000	11.050
NIROM	0.000	0.008	0.003200	0.011

NIROM, non-intrusive reduced-order modelling.

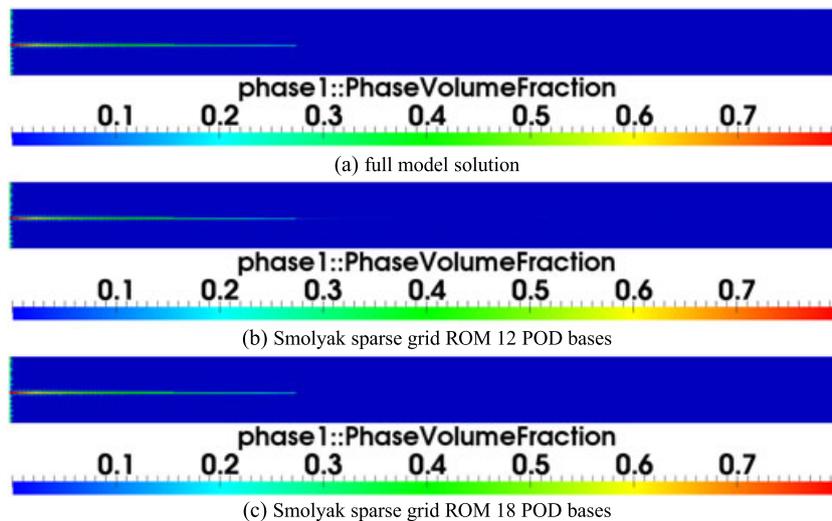


Figure 8. Water saturation distribution of the fracture problem at times 0.1. The solutions compare the predictions from NIROM based on POD and Smolyak sparse grid using 12 and 18 POD bases functions with full model.

because it does not involve assembling and solving the matrices process, thus resulting in a speed-up of computational cost by three orders of magnitude.

5.3. Case 3: High-permeability, high-aspect ratio domain embedded in a low-permeability domain

The third case is comprised of a high-permeability and high-aspect ratio domain embedded in a low-permeability domain. The high-permeability domain has a permeability 100 times bigger than the low-permeability domain. A thin, high-permeability domain is ubiquitous in reservoirs, and this represents a thin high-permeability layers or a discontinuous fracture. However, it is not easy to model it because of its small size and high-aspect ratio compared with a low-permeability domain.

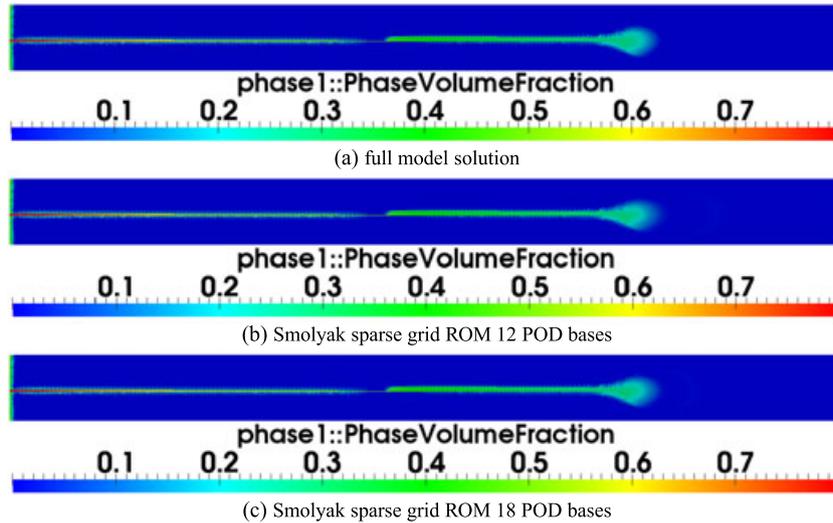


Figure 9. Water saturation distribution predicted in the fracture problem at times 0.4. The solutions compare the predictions from NIROM based on POD and Smolyak sparse grid using 12 and 18 POD bases functions with full model.

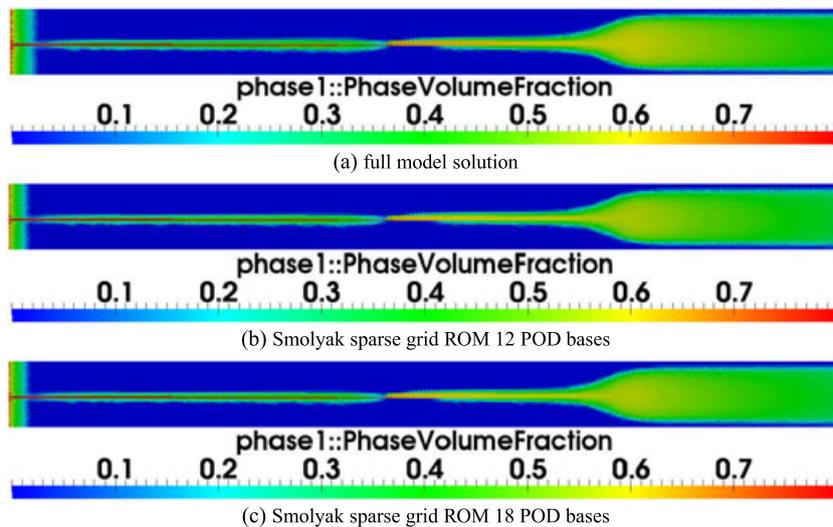


Figure 10. Water saturation distribution of the fracture problem at times 3.2. The solutions compare the predictions from NIROM based on POD and Smolyak sparse grid using 12 and 18 POD bases functions with full model.

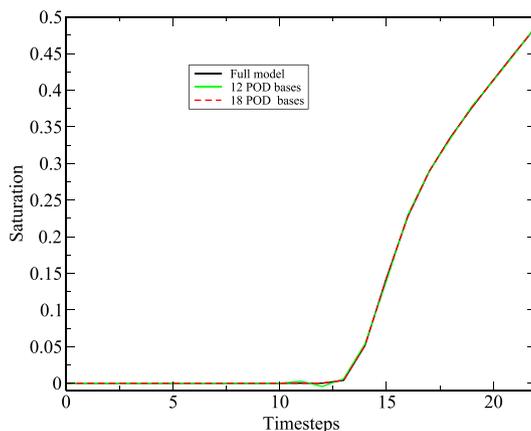


Figure 11. The saturation profile predicted by the full model and NIROM at the point $(x = 2.4671, y = 0.075)$

From the high-fidelity model simulation, with a mesh of 21 426 nodes, 25 snapshots were obtained by running full order model at regularly spaced time intervals $\Delta t = 0.14$ for each solution variables. The simulation period is $[0 - 3.5]$, and a time step of $\Delta t = 0.14$ is used for all models.

Figures 8, 9 and 10 show the solutions of saturation at time instances 0.1, 0.4 and 3.2, respectively. The solutions compare the predictions from the NIROM based on POD and Smolyak sparse grid using 12 and 18 POD bases functions with high-fidelity model. As we can see, solutions using 18 POD bases functions are much more closer to the full model.

Figure 11 shows the saturation profile predicted by the full model, and NIROM-based on POD and Smolyak sparse grid at the point $(x = 2.467, y = 0.075)$ along the computational domain. Again, we can see the NIROM performs very well with use of 12 and 18 POD basis functions.

6. CONCLUSIONS

A new non-intrusive reduced-order model for porous media multiphase flows has been presented. This method is based on a POD methodology, where optimal basis functions are generated through the method of snapshots. However, rather than using a standard Galerkin projection ROM approach (code intrusive), this approach is based on the Smolyak sparse grid interpolation non-intrusive method. This NIROM constructs a super-cube that replaces the governing equations within the reduced space. The benefits of the non-intrusive model reduction approach presented here is that it does not require any modifications to the multiphase porous media source code, due to the fact that they are independent of the equation of the system, it simply works from a number of snapshots obtained from the full multiphase porous media model.

The computational accuracy (error bound) between the Smolyak sparse grid and the full tensor product grid is given in the work of [47]. It relates to the dimensional size of the problem, approximation level and quadrature rules. The Smolyak sparse grid is less accurate than the full tensor grid. For the full tensor grid, the number of nodes required to be calculated is $(O^l)^d$. The dimensional size $d = mN_pN_{dim}N_{var}$, where m is the number of basis functions, N_p number of phases, N_{dim} dimensional size of the test cases and N_{var} is the number of variables to be solved. If we use 18 basis functions and choose velocity, pressure and saturation variables for our problems, then the dimensional size will be $18 \times 2 \times 2 \times 3 = 216$. If there are three nodes in each dimension, the total number of nodes will be 3^{216} . This is too computationally intensive to finish the simulation. In this work, the high-fidelity full model is used to numerically compare against the Smolyak sparse grid based NIROM. The results have been compared against a finite element unstructured mesh fluid model (Multiphase-Fluidity) by three cases: a 1D water flood in a homogeneous reservoir, a 2D water flood in a heterogeneous reservoir and in a simple model of a fracture. The results show that the level of accuracy of NIROM based on Smolyak sparse grid is promising. The selection of snapshots at appropriate time intervals can also affect the accuracy of the NIROM [48]. A number of

optimal selection methods of snapshots have been proposed [49–51]. In this work, all the snapshots produced at a relatively smaller time step were taken in order to ensure the accuracy. Although the accuracy of the NIROM based on Smolyak sparse grid is promising, the NIROM may not be recommended when the hypersurface is not smooth enough (i.e. hypersurface with singularities) [52].

As for the computational cost, the NIROM obtained a speed-up of around 1000 times whilst giving accurate solutions for porous media multiphase problems. The online computational cost only includes the interpolation cost and projecting cost from the reduced space onto the full space; therefore, it is computationally efficient. The offline cost includes the time forming the POD basis functions, obtaining a corresponding set of the function values and constructing the interpolation functions. The time required for forming the POD basis functions relates to the number of nodes on the mesh and the number of POD basis functions. It is not intensive and can be ignored.

In the future, we plan to evaluate the ability of the NIROM to predict flows in models with different material properties from those used to create the ROM. We will also investigate how to determine the optimum of snapshots needed to generate a ROM with the desired fidelity to the original model.

ACKNOWLEDGEMENTS

This work was carried out under funding from the British Petroleum (BP) Exploration, the UK's Natural Environment Research Council (projects NER/A/S/2003/00595, NE/C52101X/1 and NE/C51829X/1), the Imperial College High Performance Computing Service and the Engineering and Physical Sciences Research Council (GR/R60898, EP/I00405X/1 and EP/J002011/1). Prof. I.M.Navon acknowledges the support of NSF/CMG grant ATM-0931198. Dr. Xiao acknowledges the support of Janet Watson scholarship at Imperial College, NSFC grant (11502241) and China postdoctoral science foundation grant (2014M562087). Dr. Lin acknowledges the support of Imperial College-Zhejiang University Joint Applied Data Science Lab and NSFC grant (11201419). Prof. Pain and Dr. Salinas would like to thank EXXON for helping to fund the new unstructured mesh model used here; Prof. Pain, Prof. Muggeridge and Dr. Fang are grateful for the support of BP Exploration on reduced-order modelling under the LAMBSID project. Prof. Pain is grateful for the support of the EPSRC MEMPHIS multiphase flow programme grant. Prof. Muggeridge's Chair is part funded by Total.

REFERENCES

1. Batycky R, Blunt MJ, Thiele, MR, *et al.* A 3D field-scale streamline-based reservoir simulator. *SPE Reservoir Engineering* 1997; **12**(04):246–254.
2. Wilson KC, Durlofsky LJ. Optimization of shale gas field development using direct search techniques and reduced-physics models. *Journal of Petroleum Science and Engineering* 2013; **108**:304–315.
3. Durlofsky LJ, Chen Y. Uncertainty quantification for subsurface flow problems using coarse-scale models. In *Numerical analysis of multiscale problems*. Springer: Springer-Verlag Berlin Heidelberg, 2012; 163–202.
4. Jenny P, Lee H, Tchelepi H. Adaptive multiscale finite-volume method for multiphase flow and transport in porous media. *Multiscale Modeling & Simulation* 2005; **3**(1):50–64.
5. Jenny P, Lee SH, Tchelepi HA. Adaptive fully implicit multi-scale finite-volume method for multi-phase flow and transport in heterogeneous porous media. *Journal of Computational Physics* 2006; **217**(2):627–641.
6. He J. Reduced-order modeling for oil-water and compositional systems, with application to data assimilation and production optimization. *PhD Thesis*, Stanford University, 2013.
7. Ginevra Di Donato S, Wenfen Huang S, Martin Blunt S. Streamline-based dual porosity simulation of fractured reservoirs, 2003. Exhibition held in Denver, Colorado, U.S.A., 5 – 8 October 2003.
8. Siavashi M, Blunt MJ, Raisee M, Pourafshary P. Three-dimensional streamline-based simulation of non-isothermal two-phase flow in heterogeneous porous media. *Computers & Fluids* 2014; **103**:116–131.
9. Milliken W, Emanuel A, Chakravarty, A, *et al.* Applications of 3D streamline simulation to assist history matching. *SPE Reservoir Evaluation & Engineering* 2001; **4**(06):502–508.
10. Barker JW, Thibeau, S, *et al.* A critical review of the use of pseudorelative permeabilities for upscaling. *SPE Reservoir Engineering* 1997; **12**(02):138–143.
11. Barker JW, Dupouy P. An analysis of dynamic pseudo-relative permeability methods for oil-water flows. *Petroleum Geoscience* 1999; **5**(4):385–394.
12. Renard P, De Marsily G. Calculating equivalent permeability: a review. *Advances in Water Resources* 1997; **20**(5):253–278.
13. Tchelepi HA, Jenny P, Lee SH, Wolfsteiner, C, *et al.* Adaptive multiscale finite-volume framework for reservoir simulation. *SPE Journal* 2007; **12**(02):188–195.

14. Martini I, Rozza G, Haasdonk B. Reduced basis approximation and a-posteriori error estimation for the coupled stokes-darcy system. *Advances in Computational Mathematics* 2015; **41**(5):1131–1157.
15. Bistrián D, Navon IM. An improved algorithm for the shallow water equations model reduction: Dynamic Mode Decomposition vs POD. *International Journal for Numerical Methods in Fluids* 2015; **78**:552–580.
16. Daescu D, Navon IM. A dual-weighted approach to order reduction in 4d-var data assimilation. *Monthly Weather Review* 2008; **136**(3):1026–1041.
17. Stefanescu R, Sandu A, Navon IM. Comparison of POD reduced order strategies for the nonlinear 2D shallow water equations. *International Journal for Numerical Methods in Fluids* 2014; **76**(8):497–521.
18. Fang F, Pain C, Navon IM, Elsheikh A, Du J, Xiao D. Non-linear Petrov–Galerkin methods for reduced order hyperbolic equations and discontinuous finite element methods. *Journal of Computational Physics* 2013; **234**:540–559.
19. Xiao D, Fang F, Du J, Pain CC, Navon IM, Buchan AG, ElSheikh A, Hu G. Non-linear Petrov–Galerkin methods for reduced order modelling of the Navier–Stokes equations using a mixed finite element pair. *Computer Methods in Applied Mechanics and Engineering* 2013; **255**:147–157.
20. Xiao D, Fang F, Buchan AG, Pain CC, Navon IM, Du J, Hu G. Non-linear model reduction for the Navier–Stokes equations using Residual DEIM method. *Journal of Computational Physics* 2014; **263**:1–18.
21. Xiao D, Fang F, Pain C, Navon IM. Towards non-intrusive reduced order 3D free surface flow modelling. *Ocean Engineering* 2016.
22. Lin Z, Xiao D, Fang F, Pain C, Navon IM. Non-intrusive reduced order modelling with least squares fitting on a sparse grid. *International Journal for Numerical Methods in Fluids*, accepted with minor revision 2016.
23. San O, Borggaard J. Principal interval decomposition framework for pod reduced-order modeling of convective boussinesq flows. *International Journal for Numerical Methods in Fluids* 2015; **78**(1):37–62.
24. Wang M, Dutta D, Kim K, Brigham JC. A computationally efficient approach for inverse material characterization combining gappy {POD} with direct inversion. *Computer Methods in Applied Mechanics and Engineering* 2015; **286**:373–393.
25. Stefanescu R, Navon IM. POD/DEIM nonlinear model order reduction of an adi implicit shallow water equations model. *Journal of Computational Physics* 2013; **237**:95–114.
26. Giere S, Iliescu T, John V, Wells D. {SUPG} reduced order models for convection-dominated convectiondiffusion-reaction equations. *Computer Methods in Applied Mechanics and Engineering* 2015; **289**:454–474.
27. Corigliano A, Dossi M, Mariani S. Model order reduction and domain decomposition strategies for the solution of the dynamic elasticplastic structural problem. *Computer Methods in Applied Mechanics and Engineering* 2015; **290**:127–155.
28. Diez M, Campana EF, Stern F. Design-space dimensionality reduction in shape optimization by karhunenlove expansion. *Computer Methods in Applied Mechanics and Engineering* 2015; **283**:1525–1544.
29. Hoang K, Fu Y, Song J. An hp-proper orthogonal decompositionmoving least squares approach for molecular dynamics simulation. *Computer Methods in Applied Mechanics and Engineering* 2016; **298**:548–575.
30. Xiao D, Yang P, Fang F, Xiang J, Pain C, Navon IM. Non-intrusive reduced order modelling of fluid–structure interactions. *Computer Methods in Applied Mechanics and Engineering* 2016; **303**:35–54.
31. Alotaibi M, Calo VM, Efendiev Y, Galvis J, Ghommem M. Globallocal nonlinear model reduction for flows in heterogeneous porous media. *Computer Methods in Applied Mechanics and Engineering* 2015; **292**:122–137. Special Issue on Advances in Simulations of Subsurface Flow and Transport (Honoring Professor Mary F. Wheeler).
32. Xiao D, Fang F, Pain C, Navon IM, Salinas P, Muggerridge A. Non-intrusive model reduction for a 3D unstructured mesh control volume finite element reservoir model and its application to fluvial channels. *Computers & Geosciences* 2016.
33. Fang F, Zhang T, Pavlidis D, Pain C, Buchan A, Navon IM. Reduced order modelling of an unstructured mesh air pollution model and application in 2D/3D urban street canyons. *Atmospheric Environment* 2014; **96**:96–106.
34. Ansari E. Development of a surrogate simulator for two-phase subsurface flow simulation using trajectory piecewise linearization. *Journal of Petroleum Exploration and Production Technology* 2014; **4**(3):315–325.
35. Cardoso MA, Durlofsky LJ, et al. Use of reduced-order modeling procedures for production optimization. *SPE Journal* 2010; **15**(2):426–435.
36. Han C. Blackbox stencil interpolation method for model reduction. *Master's Thesis*, Massachusetts Institute of Technology, 2012.
37. Walton S, Hassan O, Morgan K. Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions. *Applied Mathematical Modelling* 2013; **37**(20):8930–8945.
38. Xiao D, Fang F, Pain C, Hu G. Non-intrusive reduced order modelling of the Navier–Stokes equations based on RBF interpolation. *International Journal for Numerical Methods in Fluids* 2015; **79**(11):580–595.
39. Klie H. Unlocking fast reservoir predictions via non-intrusive reduced order models. *The SPE Reservoir Simulation Symposium held in The Woodland, Texas, USA*, 2013; 1–16.
40. Smolyak SA. Quadrature and interpolation formulas for tensor products of certain classes of functions 1963; **4**(240-243):123.
41. Xiao D, Fang F, Buchan A, Pain C, Navon IM, Muggerridge A. Non-intrusive reduced order modelling of the Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering* 2015; **293**:552–541.
42. Pain C, Piggott M, Goddard A, Fang F, Gorman G, Marshall D, Eaton M, Power P, De Oliveira C. Three-dimensional unstructured mesh ocean modelling. *Ocean Modelling* 2005; **10**(1):5–33.
43. Bear J. *Dynamics of fluids in porous media*. Courier Dover Publications: New York US, 2013.

44. Erhel J, Mghazli Z, Oumouni M. An adaptive sparse grid method for elliptic pdes with stochastic coefficients. *Computer Methods in Applied Mechanics and Engineering* 2015; **297**:392–407.
45. Chen P, Quarteroni A, Rozza G. Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations. *Numerische Mathematik* 2015; **130**:1–36.
46. Buckley SE, Leverett, M, *et al.* Mechanism of fluid displacement in sands. *Transactions of the AIME* 1942; **146**(01):107–116.
47. Barthelmann V, Novak E, Ritter K. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics* 2000; **12**(4):273–288.
48. Siade AJ, Putti M, Yeh WW-G. Snapshot selection for groundwater model reduction using proper orthogonal decomposition. *Water Resources Research* 2010; **46**(8). doi: 10.1029/2009WR008792, 2010.
49. Lass O, Volkwein S. Adaptive POD basis computation for parametrized nonlinear systems using optimal snapshot location. *Computational Optimization and Applications* 2014; **58**(3):645–677.
50. Kunisch K, Volkwein S. Optimal snapshot location for computing POD basis functions. *ESAIM. Mathematical Modelling and Numerical Analysis* 2010; **44**(3):509–529.
51. Chen X, Navon IM, Fang F. A dual-weighted trust-region adaptive POD 4D-VAR applied to a finite-element shallow-water equations model. *International Journal for Numerical Methods in Fluids* 2011; **65**(5):520–541.
52. Geletu A. Orthogonal polynomials, quadratures & sparse-grid methods for probability integrals, 2010. A tutorial on sparse integration. Technische Universität Ilmenau, Institut für Automatisierungs- und Systemtechnik Fachgebiet Simulation und Optimale Prozesse.