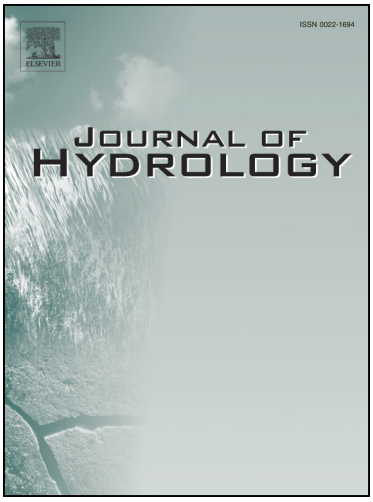
Accepted Manuscript



Model Reduction of A Coupled Numerical Model Using Proper Orthogonal Decomposition

Xinya Li, Xiao Chen, Bill X. Hu, I. Michael Navon

PII: S0022-1694(13)00656-2
DOI: http://dx.doi.org/10.1016/j.jhydrol.2013.09.011
Reference: HYDROL 19091

To appear in: *Journal of Hydrology*

Received Date: 29 January 2013
Revised Date: 7 September 2013
Accepted Date: 11 September 2013

Please cite this article as: Li, X., Chen, X., Hu, B.X., Michael Navon, I., Model Reduction of A Coupled Numerical Model Using Proper Orthogonal Decomposition, *Journal of Hydrology* (2013), doi: http://dx.doi.org/10.1016/j.jhydrol.2013.09.011

Model Reduction of A Coupled Numerical Model Using Proper Orthogonal Decomposition

Xinya Li[1], Xiao Chen[2,*], Bill X. Hu[3], and I. Michael Navon[4]

[1]Hydrology,Energy & Environment Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, United States

[2]Center for Applied Scientific Computing,Lawrence Livermore National Laboratory, Livermore, CA 94551, United States

[3]Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, FL 32306, United States

[4]Department of Scientific Computing, Florida State University, Tallahassee, FL 32306, United States

August 31th, 2013

Revised Manuscript submitted to

Journal of Hydrology

***Corresponding Author**: Tel: 925-422-6037; Email:chen73@llnl.gov

Abstract

Numerical models for variable-density flow and solute transport (VDFST) are widely used to simulate seawater intrusion and related problems. The mathematical model for VDFST is a coupled nonlinear dynamical system, so the numerical discretizations in time and space are usually required to be as fine as possible. As a result, fine-scale transient models requirelarge computational time, which is a disadvantage for state estimation, forward prediction or model inversion. The purpose of this research is to develop mathematical and numerical methods to simulate VDFST via a model order reduction technique called Proper Orthogonal Decomposition(POD) designed for nonlinear dynamical systems. POD was applied to extract leading "model features" (basis functions) through singular value decomposition (SVD) from observational data or simulations(snapshots) of high-dimensional systems. These basis functions were then used in the Galerkin projection procedure that yielded low-dimensional (reduced-order) models. The original full numerical models were alsodiscretized by the GalerkinFinite-Elementmethod (GFEM). The implementation of the POD reduced-order method was straightforward when applied to the full order model to the complex model. The developed GFEM-POD model was applied to solve two classic VDFST cases, the Henry problem and the Elder problem, in order to investigate the accuracy and efficiency of the POD model reduction method. Once the snapshots from full model results are obtained, the reduced-order model can reproduce the full model results with acceptableaccuracybutwith less computational costin comparison with the full model, which is useful formodel calibration and data assimilation problems. We found that the accuracy and efficiency of the POD reduced-order model is mainly determined by the optimal selection of snapshotsand POD bases.Validation and verification experiments confirmed our POD model reduction procedure.

Keywords: model reduction, proper orthogonal decomposition, single value decomposition, Galerkin projection, variable density flow

1. Introduction

Standard spatial discretization schemes for hydrogeological models usually lead to large-size, high-dimensional, and in general, nonlinear systems ofcoupled partial differential equations. Due to limited computational and storage capabilities, model reduction techniques provide an attractive approach to approximate the large-size discretized state equations using low-dimensional model. Thus, the model reduction techniques have received significant attention in recent years. The application of model reduction techniques for subsurface flow problems has been developed, analyzed and implemented by Vermeulenand his colleagues(*Vermeulen et al., 2004a; 2004b; 2005; Vermeulen and Heemink, 2006a*). In these pioneering studies, a proposed minimization procedure results in a significant time reduction, whereas the forward original full model must be executed certain times in order to determine optimal design or the operating parameters. The model reduction procedures developed for subsurface flow applications are based on the use of proper orthogonal decomposition (POD)(*Cardoso and Durlofsky,2010*).

Lumley (*1967*)introducedPOD in the context of analysis of turbulent flow.It is a powerful and efficient method of data analysis aiming at obtaining low-dimensional approximate descriptions (reduced-order model) of high-dimensional processes(*Holmes et al., 2012*). Data analysis using POD is often conducted to extract dominant “model characters” or basis functions, from an ensemble of experimental data or detailed simulations of high-dimensional systems, for

subsequent use in the Galerkin projection procedure that yield low-dimensional models(*Chatterjee, 2000*).This model reduction technique identifies the most energetic modes in a time-dependent system, thus providing a wayto obtain a low-dimensional description of the system's dynamics (*Fang et al., 2008*).POD reduced-order approach is introduced to transform the original flow and transport equations into a reduced form that can reproduce the dominant behaviors of the original model. The basic idea is to collect an ensemble of data of state variables (hydraulic head or solute concentration)called snapshots, by running the original model, and then use SVD to create a set of basis functions that span the snapshot collection. The reduced order model can be reconstructed using the POD basis functions by solving the rsulting set of coupled ODEs.The state variable at any time and location in the domain is expressed as a linear combination of these POD basis functions and time coefficients. Afinite-element discretization method is applied to the original model to obtain a set of ordinary differential equations for the time coefficients in the linear representation (*Kunisch and Volkwein, 2002*).

POD has been introduced and applied to various nonlinear systems (*Kunisch and Volkwein, 2002; Zheng et al., 2002; Ravindran, 2002; Meyer and Matthies, 2003; Vermeulen et al., 2006b; Cao et al., 2006; Khalil et al., 2007; Fang et al., 2008; Reis and Stykel, 2007, Siade et al., 2010*) . In practice, groundwater related problems that can be solved by a single flow model are very limited. More complicated groundwater processes are involved in coupled modeling using different numerical models. Robinson et al. (*2009*) attempted a simulation on solute transport in heterogeneous porous media using model reduction techniques. POD was also applied to multiphase (oil-water) flow (*van Doren et al., 2006*). Overall, model reduction via POD procedures is still a relatively new mathematical technique in the area of hydrogeological

modeling. Its effective application to other groundwater flow and transport processes, such as the VDFST,constitutesa challenging issue.

Numerical models of VDFST are widely used to simulateseawater intrusion and submarine groundwater discharge processes (*Bear, 1999; Diersch and Kolditz, 2002; Guo and Langevin, 2002; Voss and Provost, 2002; Li et al., 2009*).In the process of high-density fluid mixing with low-density fluid in an aquifer, fresh water flow causes the redistribution of fluid density and thus the variationofsolute concentration, and conversely affectsgroundwater movement. The groundwater flow and the solute transport are coupled processes, and the governing equations for the two processes must be solved jointly.Consequently, governing equations for VDFST problems are both transient and nonlinear. The classical numerical method, Galerkin Finite ElementMethod (GFEM), is often adopted to solve the VDFST problem, converting a continuous operator problem to a discrete problem(*Segol et al., 1975; Navon, 1979; Navon and Muller, 1979*).

In a previous study, Li and Hu (*2013*) described a POD-reduced implicit finite difference (FD) scheme (FD-POD) for uncoupled transient mass transports in heterogeneous media.However, the classical Galerkin projection method originated from the finite element analysis by defining specific weighting function (test function)tobe the same as the trial function used to compute the weighted residuals. Thus, GFEM is a natural choice for POD-based Galerkin projectionas opposed to the FD method.In other words, GFEM-POD is straightforward for implementation becausetheGFEMhas a similar weighting structure for trial solution of the POD reduced-order model.In general, the accuracy and efficiency of the GFEM-POD outweigh the counterparts for FD-POD due to its internal consistency especially when dealing with complex and coupled systems.

In this study, a GFEM-POD reduced-order method was developed to transform the original VDFST model into a low-dimensional form that can approximately reproduce thehigh-dimensional high-fidelity full model simulation results, which can be further used for state estimation, forward prediction or model inversionwith reducedcomputationalcost. To the best of our knowledge, this is the first time when POD reduction method is applied to a density-dependent flow system. Two benchmark cases were used to testify the capability of this method to approximately solve density-dependent flow problems. As a boundary controlled system, the modified Henry problem was used to test the quality of the GFEM-POD model. Additionally, the GFEM-POD model was applied to another classic VDFST problem, the Elder problem, in which the calculation results are determined by coupled governing equations and not by boundary forcing. Verification and prediction tests were performed for the two problems with various permeability distributions so as to investigate the accuracy and efficiency of the POD method in approximating the density-dependent flow fields.The developed method paves the way for future study onoptimal parameter estimation for VDFST problem based on POD reduced-order modeling.

This paper is organized as follows. In section 2, the variable density flow and solute transport model is introduced and a numerical GFEM is applied to solve the mathematical model. In section 3, the model reduction method using POD to a density dependent flow approximation is developed. The method developed here is applied to two density dependent flow problems to illustrate the efficiency and accuracy of the POD method for various scenarios tested in section 4. Finally, in section 5, we provide conclusive remarks based on the numerical findings from this study.

2. Variable Density Flow and Solute Transport (VDFST) Model

2.1. Mathematical Description of Variable-Density Flow and Solute transport Problems

Using a Cartesian coordinate system with the axes of coordinates coinciding with the principal directions of an anisotropic medium, the governing equation of two-dimensional (cross-section) variable-density flow in terms of equivalent freshwater head and fluid concentration is (*Guo and Langevin, 2002*):

$$\frac{\partial}{\partial x}\left(\rho K_{fx}\frac{\mu_f}{\mu}\frac{\partial h_f}{\partial x}\right)+\frac{\partial}{\partial z}\left(\rho K_{fz}\frac{\mu_f}{\mu}\left[\frac{\partial h_f}{\partial z}+\frac{\rho-\rho_f}{\rho_f}\right]\right)=\rho_f S_s\frac{\partial h_f}{\partial t}+\theta E\frac{\partial C}{\partial t}-\rho_{ss}q_{ss} \quad (1)$$
$$x,z\in\Omega \quad 0\le t\le T$$

where $h_f[L]$ is the equivalent freshwater head, $K_f(x,z)[LT^{-1}]$ is the freshwater hydraulic conductivity tensor, $\rho\,[ML^{-3}]$ is the fluid density, $\rho_f\,[ML^{-3}]$ is the freshwater density, μ_f/μ is the ratio of freshwater and saltwater fluid viscosity and considered equal to 1, $S_s[L^{-1}]$ is specific storage, θ is the effective porosity, $\rho_{ss}[ML^{-3}]$ and $q_{ss}[T^{-1}]$ represent the source and/or sink term, and $C[ML^{-3}]$ is the fluid concentration. Eis a dimensionless constant that represents the density-coupling coefficient, where $\begin{cases}\rho=\rho_f+EC\\ \dfrac{\partial\rho}{\partial C}=E\end{cases}$ .The relationship between concentration and density is assumed to be linear. Here, Ω represents the bounded calculation spatial domain and T is the time period of calculation.Equation (1) is subject to the following initial and boundary conditions:

$$
\begin{aligned}
&h(x,z,0)=h_0(x,z) \quad (x,z)\in\Omega \\
&h(x,z,t)\big|_{s_1}=h_1(x,z,t) \quad (x,z)\in s_1 \\
&\left(\rho K_{\mathrm{fx}}\frac{\partial h_f}{\partial x}\right)n_x+\left(\rho K_{\mathrm{fz}}\left[\frac{\partial h_f}{\partial z}+\eta c\right]\right)n_z\Big|_{s_2}=\rho_q q(x,y,t) \quad (x,z)\in s_2 \qquad (2)\\
&\mathrm{s}_1\text{:Dirichlet Boundary Condition}\\
&\mathrm{s}_2\text{:Neumann Boundary Condition}
\end{aligned}
$$

A second governing equation for the two-dimensional transport of solute mass in the porousmediais (*Guo and Langevin, 2002*),

$$
\begin{aligned}
&\frac{\partial}{\partial x}\left(D_{xx}\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial z}\left(D_{zz}\frac{\partial C}{\partial z}\right)-\frac{\partial(u_xC)}{\partial x}-\frac{\partial(u_zC)}{\partial z}=\frac{\partial C}{\partial t}-\frac{q_{ss}}{\theta}C_{ss} \qquad (3)\\
&x,z\in\Omega \quad 0\le t\le T
\end{aligned}
$$

where $D\,[L^2T^{-1}]$ is the hydrodynamic dispersion coefficient, $u\,[LT^{-1}]$ is the pore velocity, and $C_{ss}\,[ML^{-3}]$ is the solute concentration of source or sinks terms.

Equation (3) is subject to the following initial and boundary conditions,

$$
\begin{aligned}
&c(x,z,0)=c_0(x,z) \quad (x,z)\in\Omega \\
&c(x,z,t)\big|_{s_1}=c_1(x,z,t) \quad (x,z)\in s_1 \\
&\left(D_{xx}\frac{\partial c}{\partial x}\right)n_x+\left(D_{zz}\frac{\partial c}{\partial z}\right)n_z\Big|_{s_2}=g(x,z,t) \quad (x,z)\in s_2
\end{aligned}
\qquad (4)
$$

Darcy's Law is adopted in the variable-density form as,

$$
\begin{aligned}
u_x&=-\frac{K_{\mathrm{fx}}}{\theta}\frac{\partial h_f}{\partial x}\\
u_z&=-\frac{K_{\mathrm{fz}}}{\theta}\left(\frac{\partial h_f}{\partial z}+\eta c\right)
\end{aligned}
\qquad (5)
$$

Inserting (5) into (1) and (3) and using the empirical linear relation between the saltwater density and concentrationwe obtain,

$$\frac{\partial}{\partial x}\left((1+\eta C)K_{fx}\frac{\partial h_f}{\partial x}\right)+\frac{\partial}{\partial z}\left((1+\eta C)K_{fz}\left[\frac{\partial h_f}{\partial z}+\eta C\right]\right)=S_s\frac{\partial h_f}{\partial t}+\theta\eta\frac{\partial C}{\partial t}-\frac{\rho_{ss}}{\rho_f}q_{ss}$$
$$\eta=\frac{E}{\rho_f},\quad \frac{\rho}{\rho_f}=1+\eta C,\quad x,z\in\Omega\quad 0\le t\le T \tag{6}$$

$$\frac{\partial}{\partial x}\left(D_{xx}\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial z}\left(D_{zz}\frac{\partial C}{\partial z}\right)+\frac{\partial}{\partial x}\left(\frac{K_{\mathrm{fx}}}{\theta}\frac{\partial h_f}{\partial x}C\right)+\frac{\partial}{\partial z}\left(\frac{K_{\mathrm{fz}}}{\theta}\left(\frac{\partial h_f}{\partial z}+\eta C\right)C\right)=\frac{\partial C}{\partial t}-\frac{q_{ss}}{\theta}C_{ss}$$
$$x,z\in\Omega\quad 0\le t\le T \tag{7}$$

Eqs. (6) and (7) are the governing equations of a coupled nonlinear system of VDFST.

2.2 Numerical GFEM Solutions

The approximate solutions for hydraulic head and solute concentration in Eq. (6) and (7) are defined in Eq. (8) using the nodal basis function according to Galerkin finite element method (*Xue and Xie, 2007*),

$$h_f(x,z,t)\approx\tilde{h}(x,z,t)=\sum_{L}^{NNODE}h_L(t)N_L(x,z)$$
$$C(x,z,t)\approx\tilde{c}(x,z,t)=\sum_{L}^{NNODE}c_L(t)N_L(x,z) \tag{8}$$

where $h_L(t)$ is the approximated hydraulic head at node L (L = 1, …,*NNODE*)and time t, $c_L(t)$ is the approximate solute concentration at node $L$and time t. $N_L(x,z)$ is the finite-element basis function, *NNODE*(or *NN*) is the total number of nodes used across the domain.

An implicit time-extrapolated method was used to integrate the resulting system of ordinary differential equationsin time. The boundary conditions must be implemented into the global matrices by modifying the global matrices in GFEM. Aquifer parameters such as hydraulic conductivity distribution in space are represented in an element-wise discrete way (*Voss and Provost, 2002*).The coupling between flow and transport is accomplishedthrough the synchronous approach(*Guo and Langevin, 2002*),iterating the solutionsbetween the flow and

transport equations.This kind of procedure leads to a larger amount of calculation effort,compared with the constant-density flow and transport model due to the additional coupling loop and also entails additional difficultites when implementing parts of the POD model. The application of POD model will significantly reduce computation time in such a calculation- intensive system.

3. Model Reduction using Proper Orthogonal Decomposition (POD)

The reduced-order model construction methodology is given in Figure 1, modified from Vermeulen et al. (*2004b*). First, the original full numerical model is run to generate several snapshots of model states. Second, we extract dominant patterns (the basis functions) from these state snapshots via SVD. These two steps can be treated as the preprocessing steps for the reduced-order model. With the unchanged numerical formulation and system inputs (e.g. parameters, boundary conditions, initial conditions) of the original model, the selected bases are used in Galerkin projection. The Galerkin projection is the central procedure used toconstructthe reduced-order model by projecting both the partial differential equations of groundwater flow and solute transport into a low-dimensional space. After the projection step, the reduced-order model is able to simulate the same model behaviors through the reconstruction of model states with a significantly reduced computational burden. In this section, we will describe the condensed formulation of the GFEM-POD model, which is capable of simulating the coupled process of VDFST.

3.1. Snapshots and SingularValue Decomposition

As known for the VDFST model, the most importantsimulation results from the numerical model as described above are the equivalent freshwater heads and the solute concentrations in the model domain. The two variables are sampled from simulation results at defined time stepsduring the simulation period referred to as snapshots. An ensemble of nodal-value represented snapshots chosen in the analysis time interval [0, T] can be written as(*Chen et al., 2011*):

$$\begin{aligned} &\left\{h_f^1, \quad h_f^2, \quad ..., \quad h_f^{ns}\right\} \quad h_f^k \in R^{NN}, \quad k = 1,2,...,ns \\ &\left\{c^1, \quad c^2, \quad ..., \quad c^{ns}\right\} \quad c^k \in R^{NN}, \quad k = 1,2,...,ns \end{aligned} \tag{9}$$

where*ns* is the number of snapshots and NN is the number of nodes across the mesh, the vectors h_f^k and c^k both have NN entries:

$$\begin{aligned} h_f^k &= \left(h_{f,1}^k \quad ... \quad h_{f,\mathrm{NN}}^k\right)^T \\ c^k &= \left(c_1^k \quad ... \quad c_{\mathrm{NN}}^k\right)^T \end{aligned} \tag{10}$$

The collection of all h_f^k results in a rectangular $NN \times ns$ matrix R_h,and the collection of all c^k results in a rectangular $NN \times ns$ matrix R_c. The aim of POD is to find a set of orthonormal basis functions of R_h and $R_c$respectively that can capturemost of energetic information in the original VDFST system(*Fang et al., 2008*).

Singular Value Decomposition (SVD) is a well-known technique for extracting dominant "features" and coherent structures from data and "compressing" that information into a few low order "weights" (singular values) and associated orthonormal eigenfunctions(*Golub and van Loan, 1996*). The SVD of the matrix *R*, is calculated through the equation,

$$R = USV^T \tag{11}$$

where U isan $NN \times NN$ orthogonal matrix whose columns are constructed by the singular eigenvectors of RR^T, V is an $ns \times ns$ orthogonal matrix whose columns are constructed by the eigenvectors of $R^T R$, and S is a diagonal $NN \times ns$ matrix with singular values. The singular values in S are square roots of the eigenvalues from RR^T or $R^T R$. The singular values are arranged in descending order.An optimal rank m approximation to R is calculated by,

$$R_m = US_m V^T \qquad (12)$$

In computation, one would actually replace U and V with the matrices of their first m columns; and replace S_m by its leading $m \times m$ principal minor, the sub-matrix consisting of first m rows and first m columns of S. The optimality of the approximation in Eq. (12) lies in the fact that no other rank m matrix can be closer to R in the Frobenius norm, which is a discrete version of the L_2 norm (*Chatterjee, 2000*). So the first m-th columns of the matrix U (for any m) give an optimal orthonormal basis for approximating the data. The basis vectors are given by:

$$\psi_i = U_i, \quad 1 \le i \le M \qquad (13)$$

where M is the number of basis functions.

SVDis applied to snapshots matrices R_h and R_c,respectively, to obtain the POD basis functions of head and concentration:

$$\begin{aligned} \Psi^h &= \left\{ \psi^{h,1}, \psi^{h,2}, \ldots, \psi^{h,M_h} \right\} \\ \Psi^c &= \left\{ \psi^{c,1}, \psi^{c,2}, \ldots, \psi^{c,M_c} \right\} \end{aligned} \qquad (14)$$

where M_h is the number of bases from snapshots of hydraulic head, M_c is the number of bases from snapshots of solute concentration.

The eigenvalues λ_i are real and positive, and they are sorted in descending order where the ith eigenvalue is a measure of the information transferred within the ith basis mode(*Fang et al.,*

2008). Hence, if λ_i decays very fast, the basis functions corresponding tosmall eigenvaluesmay be neglected. The following formula is defined as the criterion of choosing a low-dimensional basis of size *M* (M<< ns)(*Fang et al., 2008*):

$$I(M) = \frac{\sum_{i}^{M} \lambda_i}{\sum_{i}^{ns} \lambda_i} \quad (15)$$

where*I*(M) represents the percentage of informationwhich is captured by the POD basis $\Psi_1, \ldots, \Psi_m, \ldots, \Psi_M$. This equation is used for both heads and concentrations.

3.2. Generation of POD Reduced-Order Model Using Galerkin Projection

To obtain thePOD reduced-order model, we solved the numerical models of (6) and (7) to obtain an ensemble of snapshots to generate POD bases, and then used a Galerkin projection scheme to project the model equations onto the subspace spanned by the POD basis elements(*Chen et al., 2011*). The POD solution can be expressed as(*Chatterjee, 2000; Pinnau, 2008*):

$$\begin{cases} h_f^{POD}(x,z,t) = \sum_{i=1}^{M_h} \psi^{h,i}(x,z)^{FEM-POD} \alpha_i^h(t) \\ c^{POD}(x,z,t) = \sum_{i=1}^{M_c} \psi^{c,i}(x,z)^{FEM-POD} \alpha_i^c(t) \end{cases} \quad (16)$$

where $\psi^i(x,z)$ are POD basis functions, also known as POD modes. These modes can be used to incorporate characteristics of the solution into a bounded problem by using results from numerical simulation and/or observationaldata. $h_f(x,z,t)$ and $c(x,z,t)$ are decomposed into linear combinations of time coefficients and POD modes which are the functions of space.

The POD modes are interpolated using finite element basis functions to form the GFEM-POD modes as(*Aquino et al., 2009*):

$$\begin{cases} \psi^{h,i}(x,z)^{FEM-POD} = \sum_{j=1}^{NN} N_j(x,z)\psi_j^{h,i} \quad i=1,...M_h \\ \psi^{c,i}(x,z)^{FEM-POD} = \sum_{j=1}^{NN} N_j(x,z)\psi_j^{c,i} \quad i=1,...M_c \end{cases} \tag{17}$$

where $\{\psi^i\}$ is a column vector that contains the nodal values of mode *i*.

The POD involves a mathematical procedure that transforms a number of possibly correlated state variables into a smaller number of uncorrelated variables called principal components as an ensemble of distinct state variables for the POD-reduced model.Then, the nodal-value represented POD bases are interpolated by FE-bases to obtain the corresponding FE represented continuous POD bases. Therefore, we must use a Galerkin projection approach to smooth the derivatives of the modes later(*Aquino et al., 2009*).Based on Eq. (16) and (17), corresponding finite-element represented POD solution can be expressed as(*Chen et al., 2011*):

$$\begin{cases} h_f(x,z,t) \approx \hat{h}(x,z,t) = \sum_{i=1}^{M_h}\sum_{j=1}^{NN} N_j(x,z)\psi_j^{h,i}\alpha_i^h(t) \\ c(x,z,t) \approx \hat{c}(x,z,t) = \sum_{i=1}^{M_c}\sum_{j=1}^{NN} N_j(x,z)\psi_j^{c,i}\alpha_i^c(t) \end{cases} \tag{18}$$

The model states are decomposed into linear combinations of GFEM base functions, POD modes and time coefficients.

From Eqs (6) and (7), we define two residual functions,

$$\mathbf{f_1}(h_f, c, x, z, t) = \frac{\partial}{\partial x}\left((1+\eta c) K_{\mathrm{fx}} \frac{\partial h_f}{\partial x}\right) + \frac{\partial}{\partial z}\left((1+\eta c) K_{\mathrm{fz}} \left(\frac{\partial h_f}{\partial z} + \eta c\right)\right) - S_s \frac{\partial h_f}{\partial t} - \theta\eta \frac{\partial c}{\partial t} + \frac{\rho_{ss}}{\rho_0} q_{ss}$$

$$\mathbf{f_2}(h_f, c, x, z, t) = \frac{\partial}{\partial x}\left(D_{xx} \frac{\partial c}{\partial x}\right) + \frac{\partial}{\partial z}\left(D_{zz} \frac{\partial c}{\partial z}\right) + \left(\frac{K_{\mathrm{fx}}}{\theta} \frac{\partial h_f}{\partial x}\right)\frac{\partial c}{\partial x} + \left(\frac{K_{\mathrm{fz}}}{\theta}\left(\frac{\partial h_f}{\partial z} + \eta c\right)\right)\frac{\partial c}{\partial z} - \frac{q_{ss}}{\theta}(c - c_{ss}) - \frac{\partial c}{\partial t} \quad (19)$$

The Galerkin method requires the residualsto be orthogonal with respect to the basis functions.Therefore,we need to project the original high-dimensional model onto a low-dimensional subspace generated by full model snapshots(*Vermeulen et al., 2005*).

Substituting (18) into (19) and integrating with respect to the POD bases according to Galerkin method gives:

$$\left\langle \left\langle \mathbf{f_1}(\hat{h}, \hat{c}, x, z, t), N_k \right\rangle, \psi^{h,m} \right\rangle = 0 \quad k = 1, \ldots, NN; \quad m = 1, \ldots, M_h$$
$$\left\langle \left\langle \mathbf{f_2}(\hat{c}, \hat{h}, x, z, t), N_k \right\rangle, \psi^{c,m} \right\rangle = 0 \quad k = 1, \ldots, NN; \quad m = 1, \ldots, M_c \quad (20)$$

usingthe inner product

$$\langle f, g \rangle = \int_\Omega fg d\Omega$$

andL_2norm

$$\|f\| = \langle f, f \rangle^{\frac{1}{2}}$$

In the reduced-order model, equations (6) and (7) are finally changed to:

$$\left\langle \iint_\Omega \begin{bmatrix} \frac{\partial}{\partial x}\left((1+\eta\hat{c}) K_{\mathrm{fx}} \frac{\partial \hat{h}}{\partial x}\right) \\ + \frac{\partial}{\partial z}\left((1+\eta\hat{c}) K_{\mathrm{fz}} \left(\frac{\partial \hat{h}}{\partial z} + \eta\hat{c}\right)\right) \\ -S_s \frac{\partial \hat{h}}{\partial t} - \theta\eta \frac{\partial \hat{c}}{\partial t} + \frac{\rho_{ss}}{\rho_0} q_{ss} \end{bmatrix} N_k dxdz, \quad \Psi^h \right\rangle = 0 \quad (21)$$

$$\left\langle \iint_{\Omega} \begin{bmatrix} \frac{\partial}{\partial x}\left(D_{xx}\frac{\partial \hat{c}}{\partial x}\right)+\frac{\partial}{\partial z}\left(D_{zz}\frac{\partial \hat{c}}{\partial z}\right) \\ +\left(\frac{K_{\mathrm{fx}}}{\theta}\frac{\partial \hat{h}}{\partial x}\right)\frac{\partial \hat{c}}{\partial x}+\left(\frac{K_{\mathrm{fz}}}{\theta}\left(\frac{\partial \hat{h}}{\partial z}+\eta\hat{c}\right)\right)\frac{\partial \hat{c}}{\partial z} \\ -\frac{q_{ss}}{\theta}\left(\hat{c}-c_{ss}\right)-\frac{\partial \hat{c}}{\partial t} \end{bmatrix} N_k dxdz, \quad \Psi^c \right\rangle = 0 \quad (22)$$

The key ofgenerating a POD reduced-order model is to solve system of coupled ODEs of $\alpha^c(t)$ and $\alpha^h(t)$ according to Eq. (18)-(20).This key is also known as Galerkin Projection, which involves two basic steps: (a) introducing the solutions expansions into the formulation of the deterministic or stochastic or any general spectral problem and (b) project the resulting spectral system onto the basis chosen to be the same as one for spectral expansion to yield a set of ordinary differential equations that the expansion coefficients must satisfy and construct in the POD-reduced dynamical system.

The integrations in equation (21) and (22) are the same as those for the numerical full model. The trial solutions substituted into (19) are now equation (18) rather than equation (8). Finite-element basis function has a different expression for each element, so Eq. (19) must be calculated per element before making the summation of all the elements. It should be noted that the GFEM basis functions $N_j(x,z)$ are the only spatial functions related to the areal integration of each element. Since POD bases Ψ^h and Ψ^c, and time coefficients α^h and α^c are not spatial functions,they can be extracted out of the areal integrations(*Chen et al., 2011*).

The coupled system ODEs of $\alpha^c(t)$, $\alpha^h(t)$ are expressed as,

$$\begin{cases} A_1\alpha^h+\left(\alpha^c\right)^T A_2\alpha^h+A_3\alpha^c+\left(\alpha^c\right)^T A_4\alpha^c+A_5\frac{d\alpha^h}{dt}+A_6\frac{d\alpha^c}{dt}=F_1 \\ B_1\alpha^c+\left(\alpha^h\right)^T B_2\alpha^c+\left(\alpha^c\right)^T B_3\alpha^c+B_4\frac{d\alpha^c}{dt}=F_2 \end{cases} \quad (23)$$

along with the initial conditions:

$$\begin{cases} \alpha_m^h(t_0) = \left\langle h(x,z,t_0), \psi^{h,m} \right\rangle, & m = 1,\ldots,m_h \\ \alpha_m^c(t_0) = \left\langle \mathrm{c}(x,z,t_0), \psi^{c,m} \right\rangle, & m = 1,\ldots,m_c \end{cases} \tag{24}$$

where

$$\alpha^h(t) = \left(\alpha_1(t),\ldots,\alpha_{m_h}(t)\right)^T; \qquad \alpha^c(t) = \left(\alpha_1(t),\ldots,\alpha_{m_c}(t)\right)^T$$

with the matrix notation:

$$i = 1,\ldots,NN \qquad j = 1,\ldots,NN$$

$$A_1 = \left(\Psi^h\right)^T a_1\, \Psi^h; \quad a_{1i,j} = \sum_e \left\{ \iint_e \left[K_{\mathrm{fx}}{}^e \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + K_{\mathrm{fz}} \frac{\partial N_i}{\partial z}\frac{\partial N_j}{\partial z} \right] dxdz \right\}$$

$$A_2 = \left(\Psi^h\right)^T a_2\, \Psi^h; \quad a_{2i,j} = \sum_{m=1}^{M_c} \sum_e \left\{ \iint_e \eta \sum_{j=1}^{3} N_j \psi_j^{c,m} \begin{bmatrix} K_{\mathrm{fx}}{}^e \dfrac{\partial N_i}{\partial x}\dfrac{\partial N_j}{\partial x} \\ +K_{\mathrm{fz}}{}^e \dfrac{\partial N_i}{\partial z}\dfrac{\partial N_j}{\partial z} \end{bmatrix} dxdz \right\}$$

$$A_3 = \left(\Psi^h\right)^T a_3 \Psi^c; \quad a_{3i,j} = \sum_e \left\{ \iint_e \eta K_{\mathrm{fz}}{}^e \frac{\partial N_i}{\partial z} N_j dxdz \right\}$$

$$A_4 = \left(\Psi^h\right)^T a_4 \Psi^c; \quad a_{4i,j} = \sum_{m=1}^{M_c} \sum_e \left\{ \iint_e \eta^2 K_{\mathrm{fz}}{}^e \sum_{j=1}^{3} N_j \psi_j^{c,m} \frac{\partial N_i}{\partial z} N_j dxdz \right\}$$

$$A_5 = \left(\Psi^h\right)^T a_5 \Psi^h; \quad a_{5i,j} = \sum_e \left\{ \iint_e S_s N_i N_j dxdz \right\}$$

$$A_6 = \left(\Psi^h\right)^T a_6 \Psi^c; \quad a_{6i,j} = \sum_e \left\{ \iint_e \theta\eta N_i N_j dxdz \right\}$$

$$F_1 = \left(\Psi^h\right)^T \sum_e \left\{ \int_{s_2} \frac{\rho_q}{\rho_0} q N_i ds + \iint_e \frac{\rho_{ss}}{\rho_0} q_{ss} N_i dxdz \right\}$$

$$i=1,\ldots,NN \qquad j=1,\ldots,NN \qquad k=1,\ldots,NN$$

$$B_1=\left(\psi^c\right)^T b_1\psi^c;\quad b_{1i,j}=\sum_e\left\{\iint_e\left[D_{xx}{}^e\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x}+D_{zz}{}^e\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial z}\right]dxdz\right\}$$

$$B_2=\left(\psi^c\right)^T b_2\psi^c;\quad b_{2i,j,k}=\sum_{m=1}^{M_h}\sum_e\left\{\iint_e\left[\begin{array}{l}\dfrac{K_{\mathrm{fx}}{}^e}{\theta}\displaystyle\sum_{j=1}^{3}\frac{\partial N_j}{\partial x}\psi_j^{h,m}\cdot N_i\frac{\partial N_j}{\partial x}\\ +\dfrac{K_{\mathrm{fz}}{}^e}{\theta}\displaystyle\sum_{j=1}^{3}\frac{\partial N_j}{\partial z}\psi_j^{h,m}\cdot N_i\frac{\partial N_j}{\partial z}\end{array}\right]dxdz\right\}$$

$$B_3=\left(\psi^c\right)^T b_3\psi^c;\quad b_{3i,j,k}=\sum_{m=1}^{M_c}\sum_e\left\{\iint_e\frac{K_{\mathrm{fz}}{}^e\eta}{\theta}\sum_{j=1}^{3}N_j\psi_j^{c,m}\cdot N_i\frac{\partial N_j}{\partial z}dxdz\right\}$$

$$B_4=\left(\psi^c\right)^T b_4\psi^c;\quad b_{4i,j}=\sum_e\left\{\iint_e N_iN_jdxdz\right\}$$

$$F_2=\left(\psi^c\right)^T\cdot\sum_e\left\{\iint_e\frac{q_{ss}}{\theta}c_{ss}N_idxdz+\int_{s_2}gN_ids\right\}$$

The detailedderivation of the GFEM-POD model for a VDFST system is presented in *Li*(*2010*). The dimensions of the matrices A_1-A_6 and B_1-B_4 in Eq. (23) are now determined by the number of POD bases (NB) instead of the number of nodes (NN), where NB << NN. Thus, the dimension of the reduced-order model is much smaller than the dimension of the original full model, which will save a large amount of computational labor. Thesystem of coupled ODEs, Eq. (23), still need to be solved according to the same implicit scheme stated in section 2.1. The estimated nodal values of h_f and c in the domain at a certain time can be reconstructed through Eq. (16).

3.3. Error analysis

In this subsection, the error estimates between numerical solutionsof the original model and the reduced model based on POD bases are discussed.

Let $u_{NN}^{n}(n=1,2,...,T)$ generally refers to the solution of theoriginal full model, and $u_{NN}^{n\ *}(n=1,2,...,T)$ be the vector constituted with solutions of the reduced model. NN equals to the number of active nodesacross the discretization mesh. T represents the number of time steps.

If $n \in \{1,2,...,T\}$, the error estimates are obtained as follows (*Aquino et al., 2009; Di et al, 2011*):

$$\left\| u_{NN}^{n} - u_{NN}^{n\ *} \right\|_{L^2} \le \sqrt{\lambda_{(M_u+1)}} \qquad n \in \{1,2,...,T\} \quad (25)$$

where λ represents the set of the eigenvalues of the matrices RR^T or $R^T R$, R is the matrix of an ensemble of snapshots $\{u_{NN}^{l}\}(1 \le l \le L)$. M_u is the number of basis functions chosen in the reduced model.

Else, if $n \notin \{1,2,...,T\}$, when $t_l(1 \le l \le L)$ are uniformly chosen from $t_n(1 \le n \le N)$, and $\left\| \frac{\partial u_{NN}(\zeta_1)}{\partial t} \right\|_{L^2}$ and $\left\| \frac{\partial u_{NN}^{*}(\zeta_2)}{\partial t} \right\|_{L^2}$ are bounded (i.e., $\left\| \frac{\partial u_{NN}(\zeta_1)}{\partial t} \right\|_{L^2} \le \omega$ and $\left\| \frac{\partial u_{NN}^{*}(\zeta_2)}{\partial t} \right\|_{L^2} \le \omega$), the following error estimates exist(*Di et al, 2011*):

$$\left\| u_{NN}^{n} - u_{NN}^{n\ *} \right\|_{L^2} \le \sqrt{\lambda_{(M_n+1)}} + f(T, L, \Delta t, \omega) \qquad n \notin \{1,2,...,T\} \quad (26)$$

where$u$ can replaced by $h_f$or c in equation(25) and (26).Equation (25) indicates that the error can be controlled through optimal basis selection when the sampling time period of snapshots is the same as the simulation period (e.g. a reproduction test), but the error will be inevitably larger according to Eq. (26) when the sampling time period of snapshots is different from the simulation period (e.g. a prediction test). The error in prediction test is not bounded by the descending sorted eigenvalues because of the existence of an added error function $f(T, L, \Delta t, \omega)$.

The VDFST problems described in the present manuscript adopted an assumption of linear relationship between concentration and density. For coupled system, from equations (23), the

time-dependent coefficients to be determined, α^h and α^c, are solved by construction of two groups of POD bases for head and concentration. The dimensions of the matrices A_1-A_6 and B_1-B_4 are largely decreased. For each time step, the two ODEs are commonly solved iteratively until both of the alphas converge. Referring to the single state ODE (such as transient groundwater flow), the efficiency of the reduced model in this coupled system is more significant comparedwith the original model. Meanwhile, if the two types of POD bases are generated from insufficient head and concentration snapshots simultaneously, the errors are transmitted faster and accumulated in both α^h and α^c.

4. Numerical Application Cases: Henry Problem And Elder Problem

4.1. Henry Problem

Henry problem(*Henry, 1964*), a classic variable-density flow and solutetransport problem, is applied to test the proposed GFEM-POD model.The Henry problem has played a key role in understanding of seawater intrusion into coastal aquifers, and in benchmarking density dependent flow codes(*Abarca et al., 2007*). The problem has been studied for decades, and its importance on parametric analysis of seawater intrusion is still attracting great attention (*Sanz and Voss, 2006*).

Numerical programs were compiled by Li(*2010*)to solveVDFST models using GFEM.Toexamine the accuracy of thesenumerical programs, we used the same model inputs as Simpson and Clement(*2004*)to simulate a standard Henry problem (D_m = 1.62925m^2/d), except the time step is 1 minute and the convergence criteria is 10^{-6} kg/m^3 for the fluid concentration between consecutive iterations. The system reached a steady state after approximately 250 minutes. The concentration solutions from this numerical model are compared with the semi-

analytical results (*Simpson and Clement,2004*). The isochlors revealed anexcellent correspondence, as revealed by the fact thatboth the shape and position of the isochlorsmatched very well(*Li, 2010*).

By halving the recharge rate of freshwater (Q_{in}), a modified Henry problem (Simpson and Clement, *2004*) is simulated, to increasethe relative importance of the density-dependent effects as compared to the boundary forcing. It served as the original full model. All the other model inputs are still the same as the standard Henry problem. Meanwhile, the maximum grid Peclet number is reduced from 4.1 under the standard conditions to 2.8 for the modified conditions on this 41×21 grid(*Simpson and Clement, 2004*).Under the modified conditions, the isochlor distribution will be more diffuse, which can help alleviate potential oscillation near the top-right of the aquifer(*Segol et al., 1975*). The system required approximately 460 minutes CPU time for the solution when the change of fluid concentration is smaller than 10^{-3} kg/m^3 between two successive time steps. The CPU time required to simulate 500 minutes in MATLAB with a time step of 1 minute is approximately 1500 seconds for the original full model.

4.2. Model Reduction of the Henry Problem

To demonstrate the application of model reduction, POD method discussed in section 3 is illustrated using the modified Henry problem in various cases withdifferent combination of heterogeneity and anisotropy of the conductivity field in the aquifer.In the first case, a homogeneous and isotropic aquifer is considered for the modified Henry problem. The hydraulic conductivity K_f throughout the domain is 864 m/day.Following the same procedure, the original numerical model was used to generate snapshots.

For a prediction test, the snapshots were selected initially every 1 minute from the original model solutions of the first 100 minutes for both head and concentration. We have an ensemble of snapshots with a size of 100. Reduced model extracted a certain number of POD bases from the 100 snapshots to predict the head and concentration distributions in a time period of 400 minutes, from t = 101 minute to t = 500 minutes and the predicted time step is 1 minute.

The number of POD bases (NB), snapshots selection, and the predicted time length are the most important factors in this study to determine the accuracy and efficiency of the reduced model. The impactsof these three factors on prediction were investigated as follows according to the prediction test.

4.2.1. Basis selection

Previously discussed in section 3.1, in many cases, the first few eigenvalues comprise most of the total information of a matrix. Under this condition, weneed to choose an adequatenumber of bases to capture the most information to predict the concentrationwith limited calculation. The relationship between the percentage of the total information and the number of eigenvalues is illustrated in Figure 2. By retaining only the first 5 eigenvalues (NB = 5) of the ensemble of snapshots of head solutions, 99.99% of total information is extracted. However, for concentration solutions,we need more than 12 eigenvalues of the same size of snapshotsto reach the same level of percentage. Hence, concentration can be approximated and predicted from the reduced model using a number of bases exceeding 12 in order to obtain an accurate reproduction of original model.

To investigate the effect of NB on the solution accuracy, we vary the size of NB, but keep the size of the ensemble of snapshots to be 100 and the predicted time steps to be 400. The accuracy

of the computed concentrations using model reduction with various NBs is presented in Figure 3. Two error metrics are employed to compare the predicted results between the reduced model and the original full model,by calculating root mean square error (RMSE) and the correlation coefficient for each predicted time step over the domain. Correlation is defined asthe correlation coefficient of solution vector from the original model and solution vector from the reduced model, which is calculated from:

$$r = cov\left(u^{Original}, u^{Reduced}\right);\ R = \frac{r(1,2)}{\sqrt{r(1,1)r(2,2)}} \quad (27)$$

where$u$ indicates the freshwater head or the concentration.

From Figure 3, the accuracy of the reduced model is positively correlated with the number of bases. The computation time of the reduced model with different NB is listed in Table 1. As NB increasing, the required computation time increases. An optimal value of NB is important to increase the efficiency of reduce model without sacrifice the accuracy. Employing more bases during the reduction process will not efficiently increase the accuracy, but require more computation time. In Figure 3, the accuracy of the reduced model decreases gradually as the increase of prediction time steps. The accuracy of the reduced model is best at the time t = 100 minutes. The predicted results using 20 bases have a relatively lower accuracy at t = 500 minute (Figure 4 (b) and (d)) than at t = 200 minutes (Figure 4 (a) and (c)), although, there are still good matches between the reduced model and the full model. Thesimulation of reduced model only took the snapshots from the first 100 minutes of model simulation. The coefficient $\alpha(t)$ is calculated in the reduced model as a function of time. Thus, calculation error accumulates as time increases. Normally, without additional information from new snapshots, the best prediction time period will be the same as that covered by the set of chosen snapshots. That is the reasonwe need to take more than 12 bases to maintain the accuracy, not dropping to a lower level (smaller

than 99%) in the future.The computation time using the original full model to predict 400 time steps is about 1150 seconds, whereas it tookonly 5 seconds of CPU time were required for the reduced model with NB = 20 to conduct the same prediction, which runs at least 230 times faster. It runs nearly 1200 times faster when NB = 5.

4.2.2. Predicted Time Length

To overcome the problem of accuracy decrease with time, the best approach is to add updated information in the prediction period. Observations will add significant amount of information to POD modes through new snapshots. Assuming that we add only one new snapshot which is obtained from the observations at the time t = 200 minutes to the old snapshots. The number of snapshots now is 101. The prediction period is still the same, from t = 101 minutes to t = 500 minutes. The updated results are shown in Figure 5. The NB used is still 20. Comparing withFigure 3, all predicted results were significantly improved. The reduced model can be calibrated with updated information from observations or new snapshots to significantly increase the accuracy. Addition of observation data will not only greatly increase the accuracy, but also leads to a better snapshots selection. It is worth mentioning that, the computational time is still the same, and it only changedslightly by increasing the number of snapshots. The computational time is mainly determined by the NB used in reduced model.

4.2.3. Snapshot selection

The ability of a reduced modelobtained from PODto accurately represent and, in practice, replace the full model is mainly based on the manner in which the full model snapshots are obtained (*Siade et al., 2010*), because both the number of snapshots and the time intervals of

sampling will affect the accuracy of the reduced model. If the snapshots did not include enough amount of information, the reduced order model will not provide accurate results no matter how many bases are used. Therefore,as shown in Figure 1, to maximize the accuracy, it is important to optimize the snapshots by the interaction between the original full model and the reduced-order model (Kunisch and Volkwein, 2010). The number of snapshots is optimal when the addition of another snapshot does not add a significant amount of information to the reduced model(*Siade et al., 2010*).

The sampling time of snapshots from solutions of original model determines the number of snapshots. If we sampled 100 time steps from the first 100 minutes, we have 100 snapshots. 50 snapshots will be taken with a sampling time step of 2 minutes, and 25 snapshots will be taken with a sampling time step of 4 minutes. The results using different number of snapshots without changing NB are shownin Figure 6. The accuracy of the reduced model is slightly changed. The correlation coefficients are still higher than 99.99%, which means all the three ensembles of snapshots captured the dominant characters of the model. A small set of snapshots is efficient for the reduced model to perform accurately.

In subsection 4.2.2, when the snapshot size was changed because of new information was included, selection of snapshots can be reevaluated. Figure 5 showed that the accuracy is further enhanced with a selection of 101 snapshots. The importance of this new snapshot is obvious. A large number of the old snapshots from the past 100 minutes will be not necessary. Adopting as many snapshots as possible in a certain time period isnot equal to a high level of accuracy. It is predictable that the 100+1 snapshots can be reduced to 25+1 snapshots to produce the results without sacrificing the accuracy. The result indicates that a snapshot from a new time period contains much more information that a snapshot from an old period of time.

4.2.4. Heterogeneous Case

Hydraulic conductivity fields in natural media are commonly heterogeneous and anisotropic. Thus, it is required to test the application of POD method on a more "realistic" case with a variable conductivity field. The conductivity field will significantly affect the velocity field of the VDFST system, which controls solute advection and dispersion processes. In the case study, the variability of the conductivity field is represented by the pattern and parameter values of K_f in Eq. (6) and (7).

In this case study, all the other settings for both the full model and the reduced model are same as those in the homogeneous case. Weproposed two common heterogeneous cases, a random field and a zonal field. From the homogeneous cases, we notice that the influences of snapshots, bases and predicted period length on prediction must be considered. Under various field conditions, we will investigate whether the reduced model via POD can still carry out the results efficiently and accurately with heterogeneous porous medium.

The first caseemployed a hydraulic conductivity field generated by the geostatistical approach. Assume the K_f (hydraulic conductivity) field is heterogeneous and anisotropic, where K_f is assumed to satisfy a Gaussian distribution, $N(864, 200)$. The anisotropic ratio K_{fx} / K_{fz} is 5 all over the domain. The distribution of K_f in *x*-coordinate direction, K_{fx}, is displayed in Figure 7. The range of the parameter values is 200 m/day ~ 1400 m/day.Employing 20 bases from 100 snapshots for this case, the reduced model runs approximately 250 times faster than the full model. Comparing the predicted results (Figures8 - 9), the accuracy of the reduced model is illustrated according to the continuous good fit of head and concentration distributions with time between the full and the reduced model respectively.

The second case employed a zonal heterogeneous medium. It is assumed that the K_f field is zonally distributed and anisotropic. The anisotropic ratio K_{fx} / K_{fz} is still 5 all over the domain. The distribution of K_{fx} field is displayed in Figure 10. The confined aquifer is divided into four zones. There are two patterns adopted to present the hydraulic conductivities. In this confined aquifer whose depth is 1m, the hydraulic conductivities decrease from zone 1 to zone 4 by depth in case A, and increase by depth from zone 1 to zone 4 in case B (Figure 10).

No matter which pattern is chosen, the same procedure of model reduction is conducted. To run the reduced model efficiently while retaining calculation accuracy, 25 snapshots are sampled from the first 100 minutes, which is 1 snapshot every 4 minutes. 10 bases are then computed from SVD. The spatial and temporal distributions of head and concentration over a period of 400 minutes are then solved from the reduced model.

For case A, the computation time of the reduced model is nearly 950 times faster than the full model. Figure 11 shows the spatial distributions of hydraulic head and concentration at time t = 500 minutes, which are identical with the results from the full model.

For case B, the computation time of the reduced model is nearly 750 times faster than the full model. Figure 12 shows the spatial distributions of hydraulic head and concentration at time t = 500 minutes, which are almost perfectly matched with the results from the original full model.

4.3. Model Reduction of the Elder Problem

As a boundary controlled system, the modified Henry problem was used to study the accuracy and efficiency of the GFEM-POD reduced model in section 4.2. TheGFEM-POD reduced model is applied to another classic VDFST problem, the Elder problem. The Elder problem(*Elder, 1967a; 1967b; Voss and Souza, 1987*)described a laminar fluid flow in a closed

rectangular aquifer and is commonly used to verify variable-density groundwater codes(*Simpson and Clement, 2003*).Compared with Henry Problem, the Elder problem has the characteristicthat the calculation results are only determined bycorrectly coupled governing equations, not by boundary forcing. As a result, the Elder problem will be influenced more by nonlinearity induced by variable-density condition.

For the Elder problem, we only consider advection and diffusion without dispersion. The coupled governing equations are still Eq. (6) and (7).To amplify the change of isolines of concentration by accelerating transport process and to avoid three stable steady state solutions (van Reeuwijk et al.,2009), a modified Elder problem is taken where the molecular diffusion coefficient (D_m) was doubled. For this modified Elder problem, the domain is regularly discretized using $61\times 31=1891$ nodes and 3600 triangular elements. A uniform time interval of 5 days is used for a simulation period of 5 years.All the other settings are still same as the standard Elder problem(*Simpson and Clement, 2003*). This modified Elder problem is used as the original full model. The five-year evolution of the dense fluid in this confined aquifer is shown in Figure 13. With symmetric system settings, the distribution of the plume lobes is also symmetric along the centerline of the aquifer.

The full MATLAB code solving standard or modified Elder problem was adjusted from the code for the Henry problems. The CPU time in MATLAB to simulate 5 years with a time step of 5 days is approximately 3 hours for the original full model.

In the previous section, the reduced model is applied only to predict the results for modified Henry problems. The performance of model reduction is verified through different patterns of space variation. The importance of snapshots selection and bases selection is discussed.

To further investigate the quality of the reduced model for Elder problem, two types of calculation are performed, reproductionand the prediction. For the reproduction calculation, the simulation period of the reduced model is the same as the time period usedin the full model to generate snapshots. While for prediction calculation, the simulation period of the reduced model is beyond the time period for the full model to generate snapshots. Based on the error analysis in section 3.3, the errors of reproduction test are addressed by equation (25) and the errors of prediction test are expressed by equation (26). From the error analysis, the errors of reproduction test can be controlled through optimal snapshots selection and base selection, which determine the $(M+1)^{th}$ eigenvalue. The errors of prediction tests are not only determined by the eigenvalues, but also by selected time period length and a case-specific coefficient. It is much more difficult to control the errors for prediction tests. The accuracy will decrease gradually as the prediction time increases. Therefore, the accuracy and efficiency of the reduced model have to be discussed according to different objects of reduced modeling.

4.3.1. Reproduction Calculation

The reproduction test is the repeated calculation of the forward simulation of the full model. The original full model was operated to simulate a time period of five years (1825 days) with a uniform time interval of 5 days. 73 snapshots were chosen from the full model results for hydraulic heads and concentrations, respectively. These 73 snapshots were sampled regularly, one from every 25 days. From SVD process, 11 POD bases are selected for the reduced model, which will reproduce the same time period with a time interval of 5 days and thus using 365 time steps. The reduced model ran approximately 2500 time faster in MATLAB than the original full

model. The comparison of the dense fluid distribution is shown in Figure 13 at the end of the first year, the third year and the fifth year, respectively.

The accuracy of the reduced model is satisfied according to Figure 13. The results of the reduced model were over 99.9% matched with the results from the full model. For reproduction test, the error can be very low because the important system information in this time period is all available through optimal selection of snapshots. As long as the snapshots cover most information, the reduced model can reproduce the head and concentration results at any time inside this time period very accurately. The reproduction tests confirmed that the reduced model can be used to replace the full numerical model for state estimation and inverse modeling which normally require repeated forward run of the full model.

4.3.2 Prediction Calculation

The snapshots for prediction tests were sampled from the full-model results of first year. For the first 365 days, we selected one snapshot from each 5 days. 11 bases were selected from the 73 snapshots. We used the information from the first year to predict the results in the next two years. The time interval used in the prediction test is 5 days. The correlation of predicted concentrations for the following two years between the reduced model and the full model is shown in Figure 14. The accuracy of the reduced model decreases rapidly with increase of prediction time. At the end of the second year ((number of time steps = 146), the accuracy is nearly 99%. However, at the end of the third year (number of time steps = 219), the accuracy is only 80%. Apparently, the reduced model cannot attain a satisfactory predictionin a time period longer than one year for this modified Elder problem, if the accuracy must be kept higher than 99% by a modeler.

More snapshots were included and more basis functions were adopted trying to predict more accurate results. However, the precision of the predicted results at the end of the third year is still not satisfied. As mentioned previously, the errors generated in prediction calculation will increase inevitably as the increase of predicted time length. The errors cannot be reduced by choosing more POD bases produced from the unchanged ensemble of snapshots. Elder problem is much more dependent on the accuracy of the coupling scheme. The evolution profile of the lobes and the fingering pattern is significantly different from previous time period, thus the snapshots extracted from a previous duration lose effectiveness of contained information rapidly during the state estimation of the predicted interval.

In section 4.2.2, we proposed an appropriate approach to overcome the problem of accuracy decrease with time, adding updated information in the prediction period. The principle is very similar to the process of weather forecasting. The reduced model is kept running, but the snapshots used also need to be updated. Observations at a certain time in the prediction period will add significant amount of new information. Illustrated by Figure 5, new snapshots are obtainedfrom observations and are added to the old ensemble of snapshots. The updated snapshots are then applied in the reduced model to increase model prediction accuracy. This updating is continuously conducted to maintain the accuracy of the reduced model.

To investigate efficiency of this method, another case is designed. The concentration results of the reduced model from the previous prediction test are compared with the results of the full model (Figure 15, (a) and (b)) at the end of the 2nd year. The snapshots are all sampled from the first year. Although, the two contours display a good fitting with each other, the transport depths of the lobes at both sides do not match well, which is marked by the red dashed line in Figure 15. It is assumed that weobtained a small set of observation data at a certain time point early in the

2nd year which was imitated from the simulation of the original full model. A new snapshot is generated based on the observation data and is included it into the old snapshots. With updated snapshots, we reran the reduced model to predict results in the same time period. The simulation results are clearly improved (Figure 15, (c)).

The importance of updating snapshots indicates again that the accuracy of reduced model relies on the time period in which full-model snapshots are sampled as discussed in section 3.3. In practice, the observations need to be filtered and weighted before they are adopted in the reduced model (*Siade et al., 2010*).

5. Conclusion

In this study, we developed a POD approach toefficiently simulate a coupled nonlinear subsurface flow and transport process. An integrated methodology of model reduction was developed through combining POD with the GFEM, so it is referred to as GFEM-POD method. The GFEM-POD method can reduce the dimension of stiffnessmatrices and forcing vectors in the full finite element numerical model to a very small size. The reduced dimension depends on the selected number of basis functions.

This method is efficient because the reduced-order model represents new states in terms of the dominant basis vectors generated by a subset of old states. The simulations of the reduced-order model must be performed in a low-dimensional space depending on the proper decomposition of model states (hydraulic head and solute concentration) in space and time.

We applied this procedure to two benchmark VDFST problems with variousscenarios. These case studies results indicate that thisGFEM-POD reduced-order model can reproduce and predict the full model results of spatial distributions for both hydraulic head and solute concentration

very accurately.The computational time required for the reduced-order model is dramatically reduced compared to the timeused in the full model simulation. The calculation accuracy depends strongly on the sampling and updating strategyof the full-model snapshots. The selected snapshots further determine how many basis functions should be used in order to achieve satisfactory results in the reduced-order model. The optimal selection of snapshots and basis functions is crucial for the application of POD and should be carefully considered due to the model's mathematical and parametric structures. We also observed that the POD approachis less robust for model prediction than for model reproduction. The reduced-order model will encounter significant calculation errors for long-term prediction. This phenomenon is more obvious when the study problem ishighly mathematically nonlinear. This requires by necessity application of interpolation methodology, such as DEIM (discrete empirical interpolation method) (Stefanescu and Navon, 2013). An effective approach of alleviating this issue is to update snapshots continuously to assimilate new information from observations or experiments.

According to error analysis, two types of cases are considered: reproduction and prediction. In previous researches, such as Vermeulen et al 2005, for a groundwater flow model, the advantage of POD in the calibration process is demonstrable, because it required the repetition (reproduction) of simulation period of the original model. The accuracy of reduced model is proven high according to their error analysis, on the premise that snapshots covered most of the information. Indeed, when the snapshots failed to capture enough information, the accuracy of the reduced model is not robust any more. The application of POD should be directed into two major directions:

1. Update snapshots by assimilating reliable information from measurements, observations under the condition that the original simulation is high-fidelity.

2. The original model has large uncertainties, which may leads to a low-fidelity simulation. The reduced model is employed as a substitute in calibration process.

We will perform further investigations in future work with the emphasis on the above two topics considering the utilization of various types of field observations for both calibration and prediction.

Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore NationalLaboratory underContractDE-AC52-07NA27344.ProfNavonacknowledges the support of NSF grant ATM-0931198.

References

Abarca, E., J. Carrera, X. Sanchez-Vila, and M. Dentz (2007),Anisotropic dispersive Henry problem, *Advances in Water Resources*, 30, 913-926.

Aquino, W., J.C. Brigham, C.J. Earls and N. Sukumar (2009), Generalized finite element method using proper orthogonal decomposition,*International Journal for Numerical Methods in Engineering*, *79*(7), 887-906.

Bear, J. (1999), Mathematical modeling of seawater intrusion, in *Seawater Intrusion into Coastal Aquifers*, edited by J. Bear, et al., pp. 127-161, Kluwer Academic Publications.

Cao, Y., J. Zhu, Z. Luo and I.M. Navon (2006), Reduced order modeling of the upper tropical Pacific ocean model using proper orthogonal decomposition, *Computers & Mathematics with Applications*, *52*(8–9), 1373–1386.

Cardoso, M. A. and L. J. Durlofsky (2010), Linearized reduced-order models for subsurface flow simulation, *Journal of Computational Physics*, *229*, 681-700.

Chatterjee, A. (2000), An introduction to the proper orthogonal decomposition, *Current Science*, *78*(7), 808-817.

Chen, X., I. M. Navon and F. Fang (2011), 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*, *65*(5), 520-541.

Di, Z., Z. Luo, Z. Xie, A. Wang and I.M. Navon (2011), An optimizing implicit difference scheme based on proper orthogonal decomposition for the two-dimensional unsaturated soil water flow equation.*International Journal for Numerical Methods in Fluids* Volume 68,pp 1324--1340 (2012)

Diersch, H. -J. G. and O. Kolditz (2002), Variable-density flow and transport in porous media:

approaches and challenges, *Advances in Water Resources*, *25*(8-12), 899–944.

Elder, J.W. (1967a), Steady free convection in a porous medium heated from below,*J Fluid Mech.*, *27*, 29–50.

Elder, J.W. (1967b),Transient convection in a porous medium,*J Fluid Mech.*, *27*, 609–623.

Fang, F., C. C. Pain, I. M. Navon, M. D. Piggott, G. J. Gorman, P. E. Farrell, P.A. Allison, and A.J. H. Goddard (2008), A PODreduced-order 4D-Var adaptive mesh ocean modeling approach, *International Journal For Numerical Methods In Fluids*, *60*(7), 709-732.

Golub, G. H. and C. F. Van Loan (1996),*Matrix Computations*, 3rd Edition, John Hopkins Univ. Press,Baltimore,Maryland.

Guo, W., and C. D. Langevin (2002), User's guide to SEAWAT: A computer program for simulation of three-dimensional variable-density ground-water flow, *U.S. Geological Survey Techniques of Water-Resources Investigations*, *Book 6*, chapter A7, 77 p.

Henry, H. R. (1964), Effects of dispersion on salt encroachment in coastal aquifers,*U.S. Geological Survey Water-Supply Paper*, *1613-C*, C71-C84.

Holmes, P., J. L. Lumley and G. Berkooz (2012), *Turbulence, Coherent Structures, Dynamical System and Symmetry*(Cambridge Monographs on Mechanics), Cambridge University Press, Cambridge, 2012.

Khalil, M., S. Adhikari, and A. Sarkat (2007), Linear system identification using proper orthogonal decomposition,*Mechanical Systems and Signal Processing*, *21*(8), 3123-3145.

Kunisch, K., and S. Volkwein (2002), Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics, *SIAM Journal on Numerical Analysis*, *40*(2), 492-515.

Kunisch, K., and S. Volkwein (2010), Optimal snapshot location for computing POD basis functions,*ESAIM:M2AN*, 44: 509-529.

Li, X., B. X. Hu, W. C. Burnett, I. R. Santos and J. P. Chanton (2009), Submarine ground water discharge driven by tidal pumping in a heterogeneous aquifer, *Ground Water*, *47*(4), 558-568.

Li, X. (2010), Model simulation and reduction of variable-density flow and solute transport using proper orthogonal decomposition, Ph.D. Thesis, Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee,Florida.

Li, X, and B. X. Hu (2013), Proper orthogonal decomposition reduced model for mass transport in heterogeneous media,*Stochastic Environmental Research and Risk Assessment*, 27(5): 1181-1191.

Lumley, J. L. (1967), in *Atmospheric turbulence and radio wave propagation*, edited by A. Yaglom and V. Tatarski, pp. 166–178, Nauka, Moscow.

Meyer, M. and H.G. Matthies (2003), Efficient model reduction in non-linear dynamics using the Karhunen-Loève expansion and dual-weighted-residual methods, *Computational Mechanics* 31: 179–191.

Navon, I. M. (1979), Finite element simulation of the shallow-water equations model on a limited-area domain, *Appl. Math. Modeling*, *3*, 337-348.

Navon, I. M. and U. Muller (1979), FESW - A finite-element FORTRAN IV program for solving the shallow-water equations,*Advances in Engineering Software*, 1, 77-86.

Pinnau, R (2008), Model reduction via proper orthogonal decomposition, in*Model Order Reduction: Theory, Research Aspects and Applications*, edited by W. H. A. Schilder and H. van der Vorst, pp. 96-109, Springer.

Ravindran, S.S (2002), Adaptive reduced-order controllers for thermal flow system using proper orthogonal decomposition,*SIAM Journal of Scientific Computing*, *23*(6), 1924–1942.

Reis,T.and T.Stykel(2007),Stability analysis and model order reduction of coupled systems, *Mathematical and Computer Modeling of Dynamic Systems*, *13*(5), 413-436.

Robinson, B.A., Z. Lu, and D. Pasqualini (2009), Simulating solute transport in porous media using model reduction techniques, submitted to*Advances in Water Resources*.

Sanz, E., and C. I. Voss (2006), Inverse modeling for seawater intrusion in coastal aquifers: Insights about parameter sensitivities, variances, correlations and estimation procedures derived from the Henry problem, *Advances in Water Resources*, *29*, 439-457.

Segol, G., G.F.Pinder and W. G. Gray (1975), A Galerkin-finite element technique for calculating the transient position of the saltwater front, *Water Resour.Res.*,*11*, 343–7.

Siade, A. J., M. Putti and W.W.-G.Yeh (2010), Snapshot selection for groundwater model reduction using proper orthogonal decomposition, *Water Resour.Res.*,*46*, W08539.

Simpson, M.J. and T.P. Clement (2003), Theoretical analysis of the worthiness of Henry and Elder problems as benchmarks of density-dependent groundwater flow models, *Advances in Water Resources*, *26*, 17-31.

Simpson, M.J. and T.P. Clement (2004),Improving the worthiness of the Henry problem as a benchmark for density-dependent groundwater flow models, *Water Resour.Res.*,*40*, W01504.

Stefanescu, R. and I. M. Navon (2013), POD/DEIM nonlinear model order reduction of an ADI implicit shallow water equations model, *Journal of Computational Physics*, 237: 95-114.

van Doren, J.F.M., R.Markovinovic and J.D.Jansen (2006), Reduced-order optimal control of water flooding using proper orthogonal decomposition, *ComputationalGeosciences*, *10*, 137–158.

vanReeuwijk, M., S.A. Mathias, C.T. Simmons, and J.D. Ward (2009), Insights from a pseudospectral approach to the Elder problem. *Water Resources Research*, 45(4), W04416.

Vermeulen, P.T. M., A. W. Heemink, and C. B. M. teStroet (2004a), Low-dimensional modeling of numerical groundwater flow, *Hydrological Process* 18(8): 1487-1504.

Vermeulen, P. T. M., A. W. Heemink, and C. B. M. teStroet (2004b), Reduced models for linear groundwater flow models using empirical orthogonal functions, *Advances in Water Resources*,*27*(1), 57-69.

Vermeulen, P. T. M., A. W. Heemink, and J.R.Valstar (2005), Inverse modeling of groundwater flow using model reduction, *Water Resour.Res.*,*41*(6), W06003.

Vermeulen, P. T. M., and A. W. Heemink (2006a), Model-reduced variational data assimilation, *Monthly Weather Review*, *134*(10), 2888-2899.

Vermeulen, P. T. M., C.B. M. teStroet, and A.W. Heemink (2006b), Model inversion of transient nonlinear groundwater flow model using model reduction, *Water Resour.Res.*,*42*(9), W09417.

Voss,C.I.and W.R. Souza(1987), Variable density flow and solute transport simulation of regional aquifers containing a narrow freshwater-saltwater transition zone, *Water Resour. Res.*,*23*(10), 1851-1866.

Voss, C. I., and A. M. Provost (2002),SUTRA: A model for saturated-unsaturated,variable-density ground-water flow with solute or energy transport,*U.S. Geological Survey Water-Resources Investigations Report 02-4231*, 290p.

Xue, Y. and C.Xie (2007), *Numerical Simulation for Groundwater*, 451 pp., Science, Beijing.

Zheng, D. and K.A.Hoo, M.J. Piovoso (2002), Low-order model identification of distributed parameter systems by a combination of singular value decomposition and the Karhunen-Loève expansion, *Industrial & Engineering Chemistry Research*,*41*(6), 1545–1556.

Tables

Table 1. Computation times of the reduced-order model for the homogeneous case with different NB to predict 400 time steps.

Computation Time (seconds)	Number of Bases (NB)
0.125	**1**
0.350	**2**
0.880	**5**
1.820	**10**
3.250	**15**
4.900	**20**

Figure Captions

Figure 1.Methodology for constructing a reduced-order model.

Figure 2. (Top) The percentage of total information of head exacted as function of number of eigenvalues for the homogeneous case; (Bottom) The percentage of total information of concentration exacted as function of number of eigenvalues for the homogeneous case.

Figure3. RMSE (Top) and correlation (Bottom) of predicted concentrations between the reduced-order model and the original full model for the homogeneous case using different number of bases from 100 snapshots.

Figure 4.Comparison of results between the reduced-order model (red dash) and the original full model (blue dash) for the homogeneous case. (a) Predicted head distribution (m) at time t = 200 minutes; (b) Predicted head distribution (m) at time t = 500 minutes; (c) Predicted concentration distribution (kg/m^3) at time t = 200 minutes; (d) Predicted concentration distribution (kg/m^3) at time t = 500 minutes.

Figure 5. RMSE of predicted concentrations between the reduced-order model and the original full model for the homogeneous case with addition of a new snapshot at t = 200 minutes (red) comparing to the previous simulation without new snapshots (black).

Figure 6. RMSE of predicted concentrations between the reduced-order model and the original full model for the homogeneous case using different number of snapshots with the same NB =20.

Figure 7. Stochastic distributed hydraulic conductivity field used in the first heterogeneous case with a Gaussian distribution, N (864, 200).

Figure 8.Comparison of results between the reduced-order model (red dash) and original full model (blue dash) for the first heterogeneous case. (a) Predicted head distribution (m) at time t = 200 minutes; (b) Predicted head distribution (m) at time t = 500 minutes.

Figure 9.Comparison of results between the reduced-order model (red dash) and original full model (blue dash) for the first heterogeneous case. (Top) Predicted concentration distribution (kg/m^3) at time t = 200 minutes; (Bottom) Predicted concentration distribution (kg/m^3) at time t =

500 minutes.

Figure 10. Diagrams display, in cross-section view, the two zonal patterns and parameter values used in the second heterogeneous case. (A) Hydraulic conductivities decrease by depth; (B) Hydraulic conductivities increase by depth.

Figure 11.Comparison of results between the reduced-order model (red dash) and original full model (blue dash) for Case A using the zonal approach. (Top) Predicted head distribution (m) at time t = 500 minutes; (Bottom) Predicted concentration distribution (kg/m^3) at time t = 500 minutes.

Figure 12.Comparison of results between the reduced-order model (red dash) and original full model (blue dash) for Case B using the zonal approach. (Top) Predicted head distribution (m) at time t = 500 minutes; (Bottom) Predicted concentration distribution (kg/m^3) at time t = 500 minutes.

Figure 13.Comparison of dense fluid distribution between the reduced-order model (right) and original full model (left) in the reproduction test. The concentration contour interval is 28 kg/m^3.

Figure 14.Correlation of predicted concentrations between the reduced-order model and the original full model in the prediction test for the next 2 years with 146 time steps.

Figure 15.Predicted dense fluid distribution of the reduced-order model (a), the original full model (b) and the updated reduced-order model (c) in the prediction test at the end of the 2nd year. The concentration contour interval is 28 kg/m^3.

Table 1

Computation Time (seconds)	Number of Bases (NB)
0.125	1
0.350	2
0.880	5
1.820	10
3.250	15
4.900	20

Figure 1

Figure 1

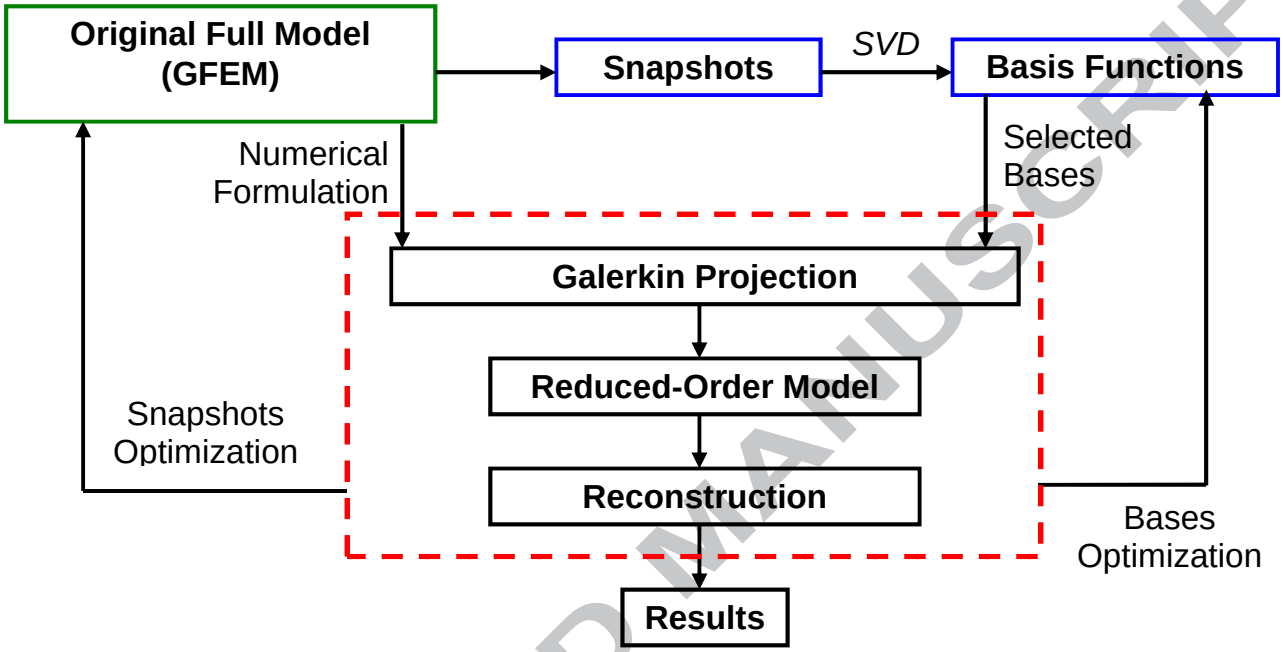
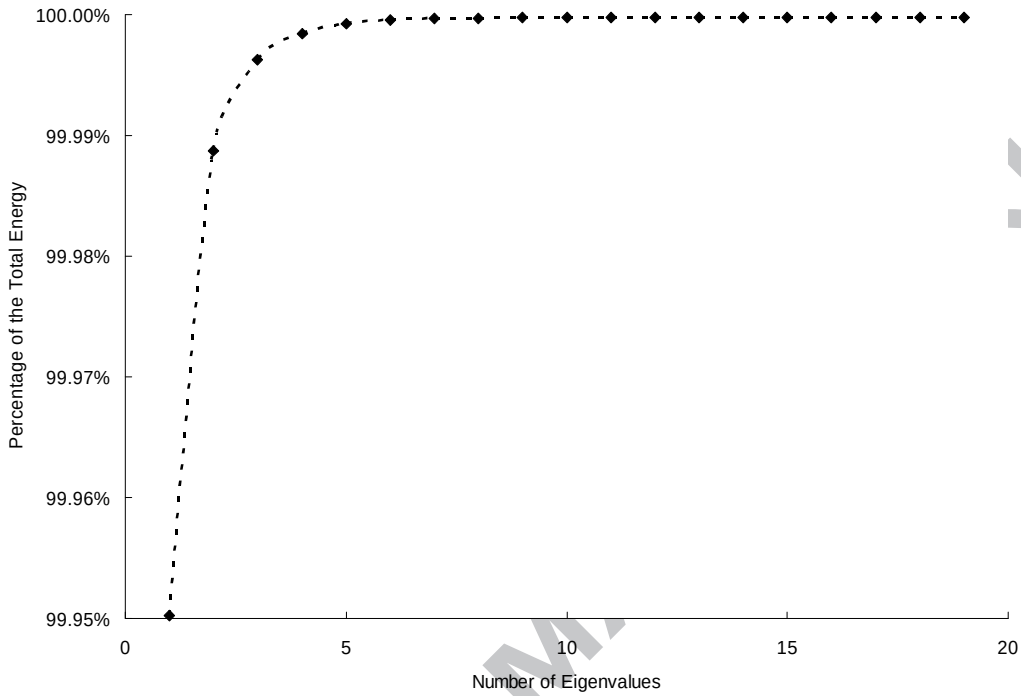


Figure 2

Figure 2



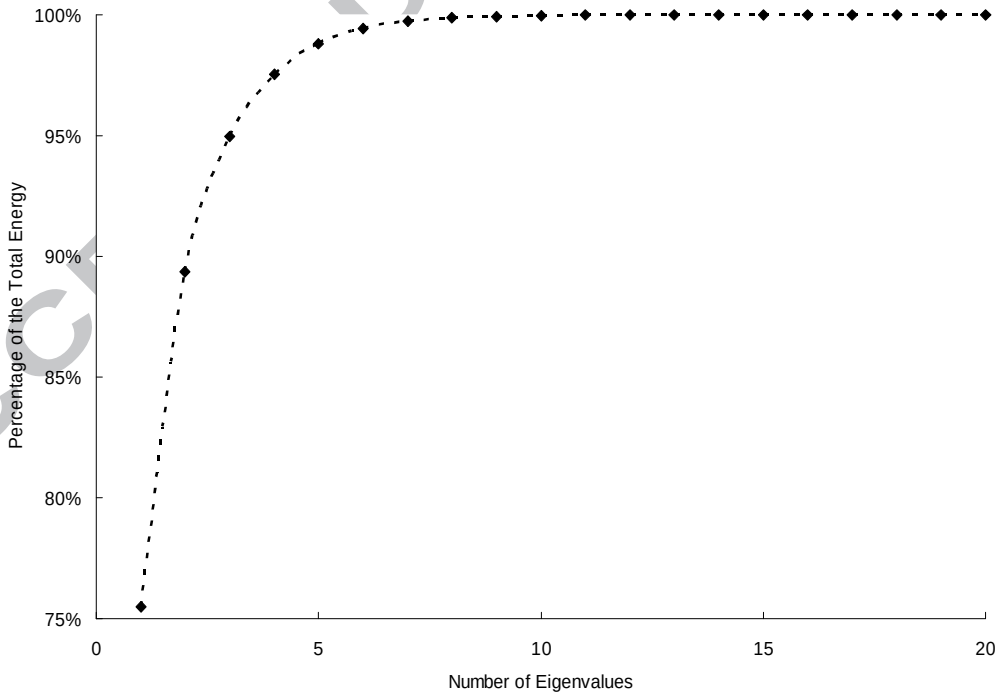
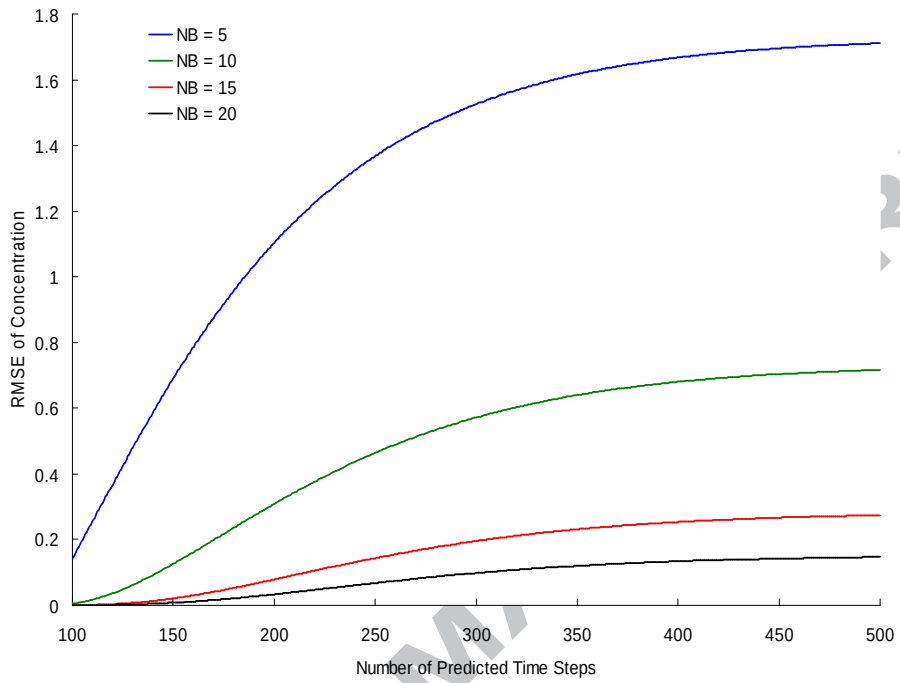


Figure 3

Figure 3



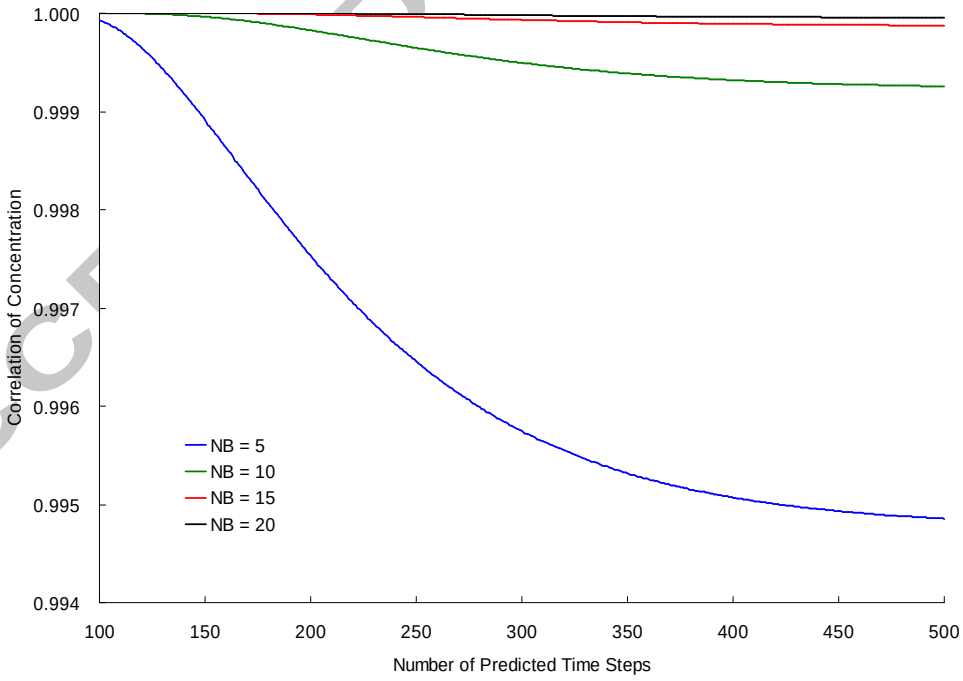


Figure 4

Figure 4

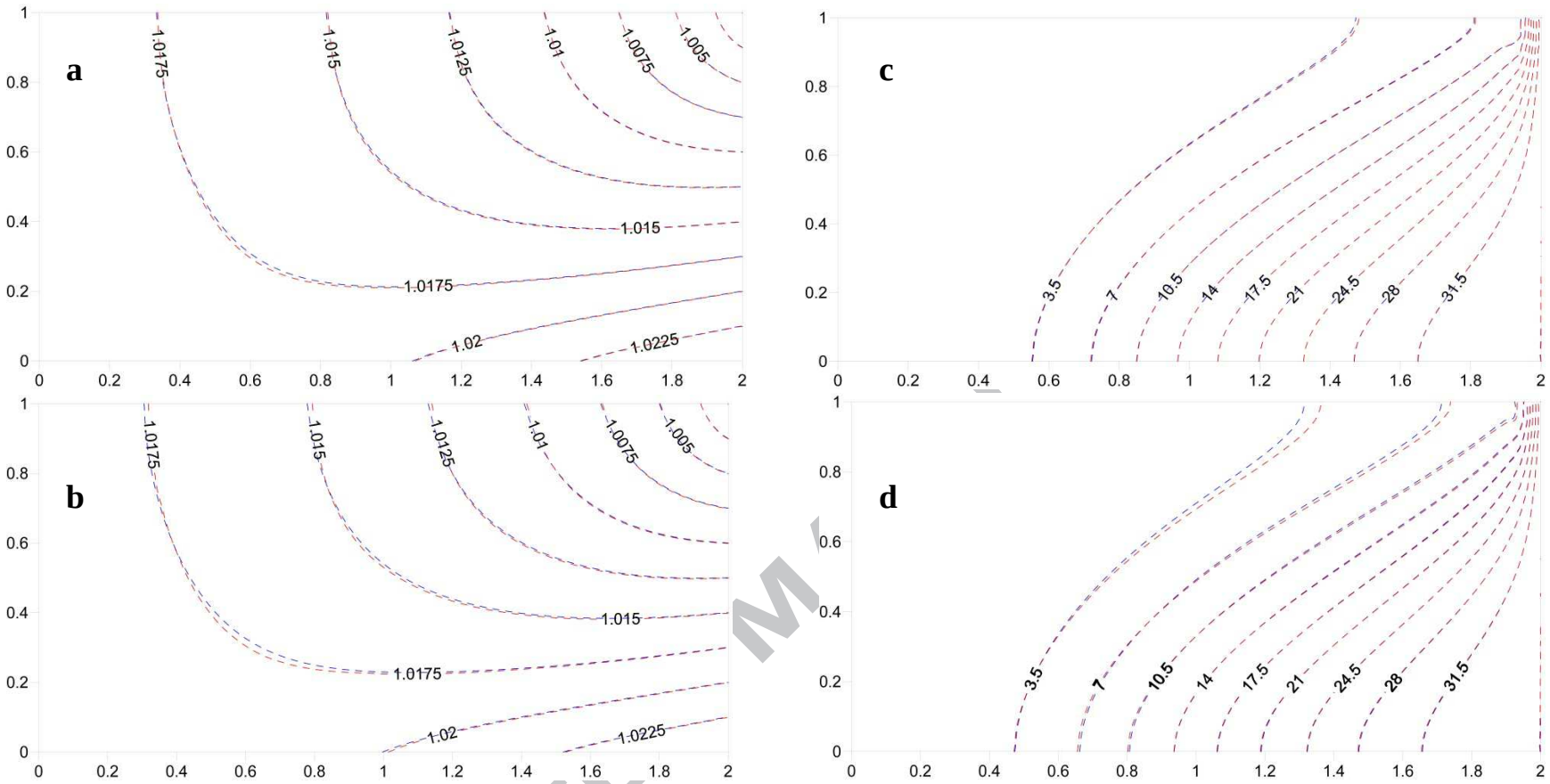


Figure 5

Figure 5

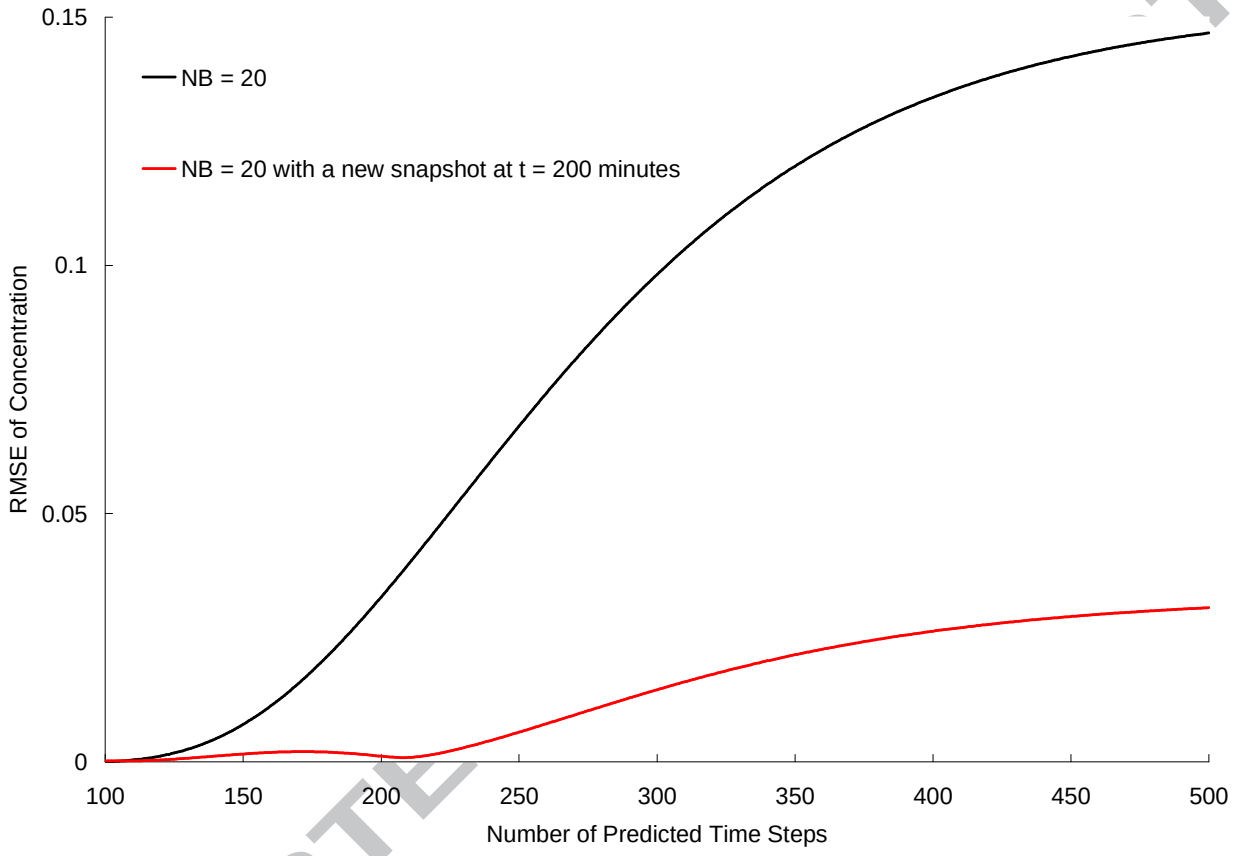


Figure 6

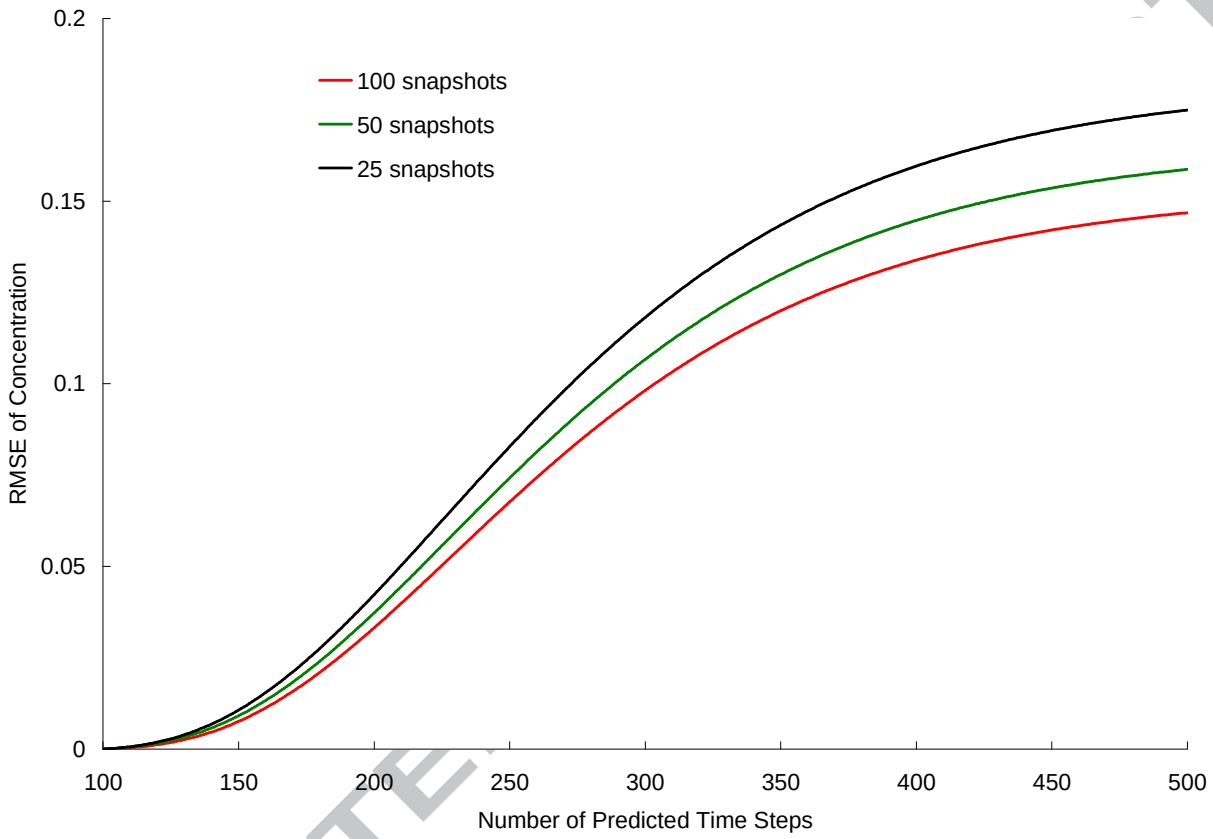


Figure 7

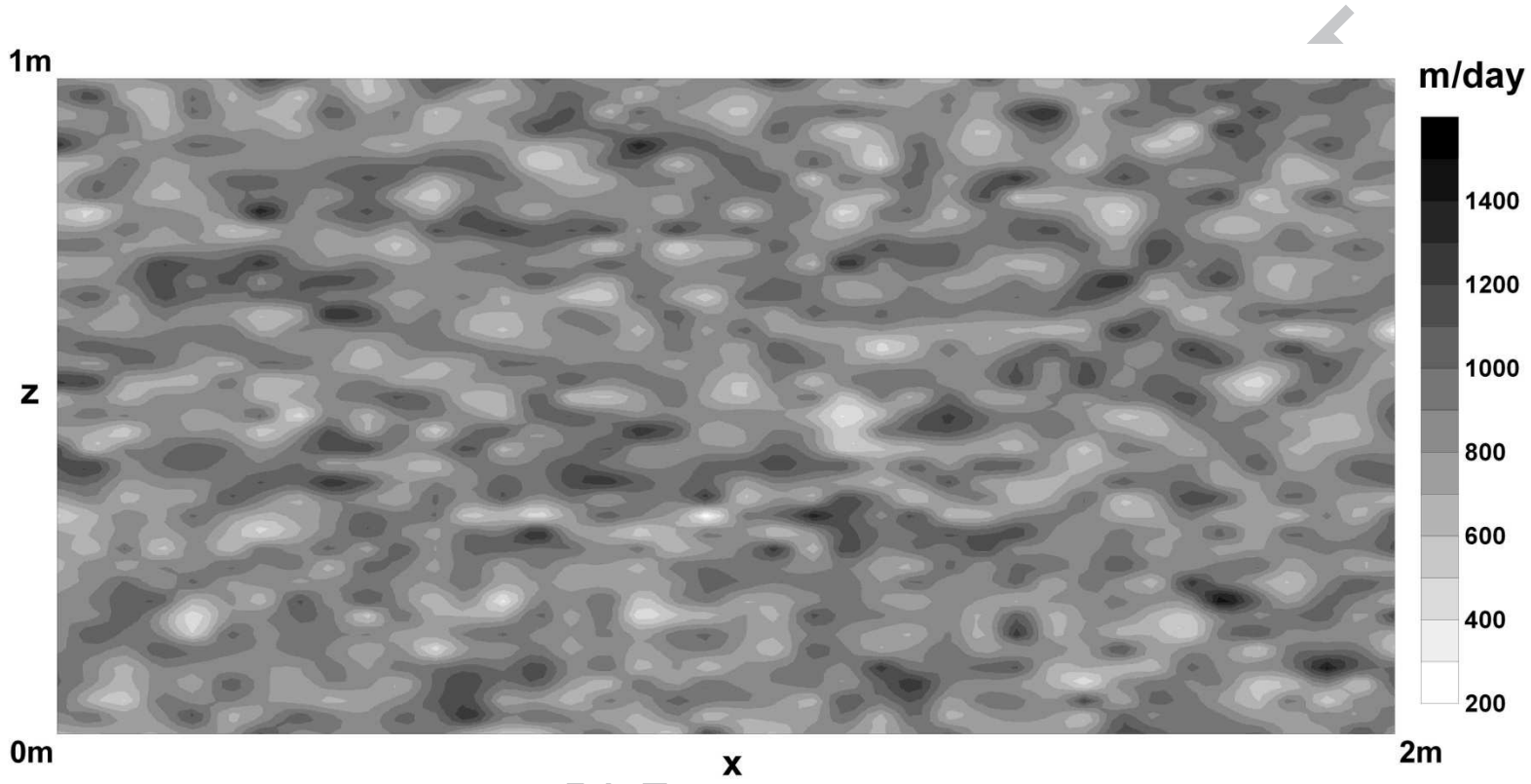


Figure 8

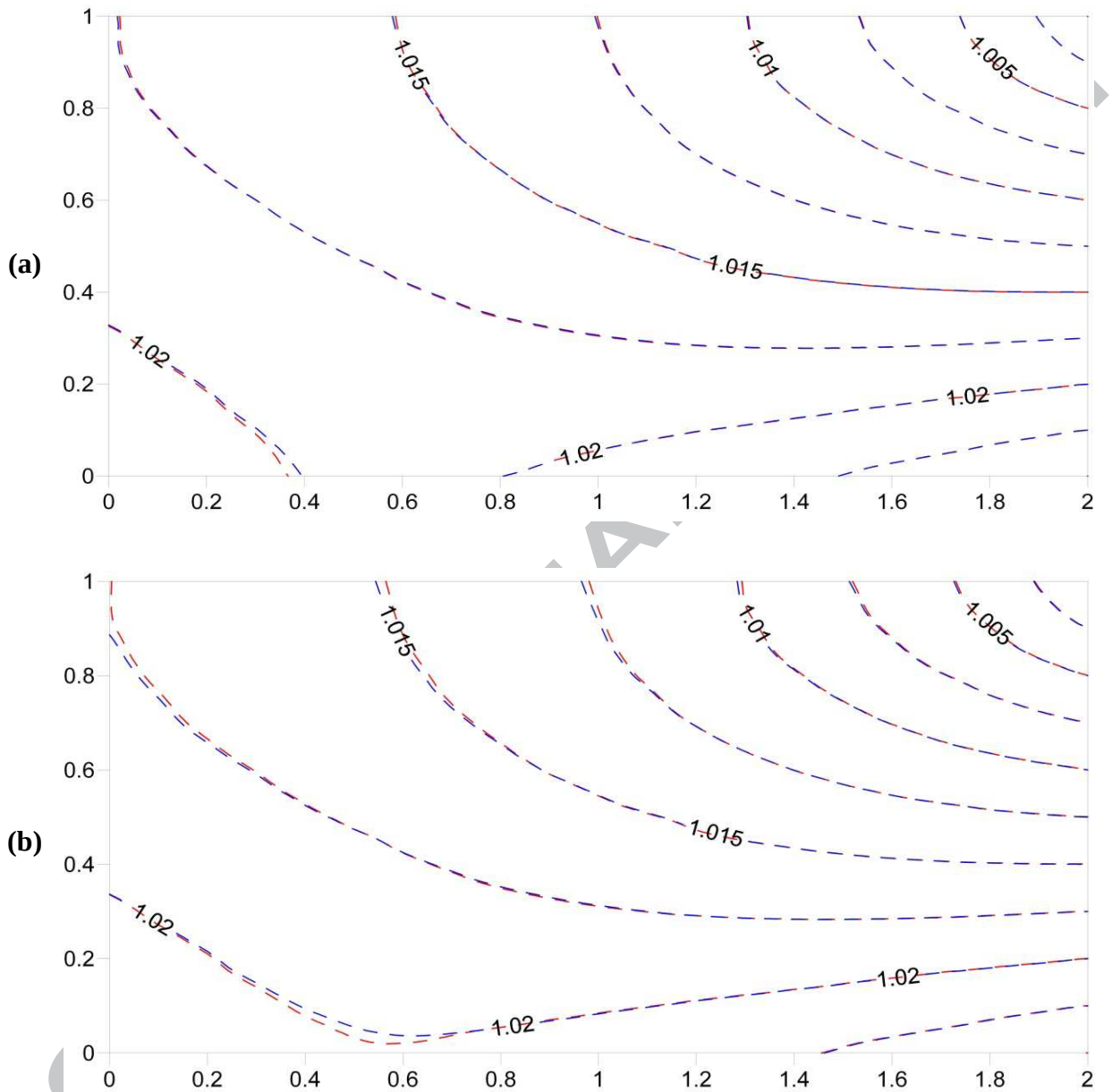
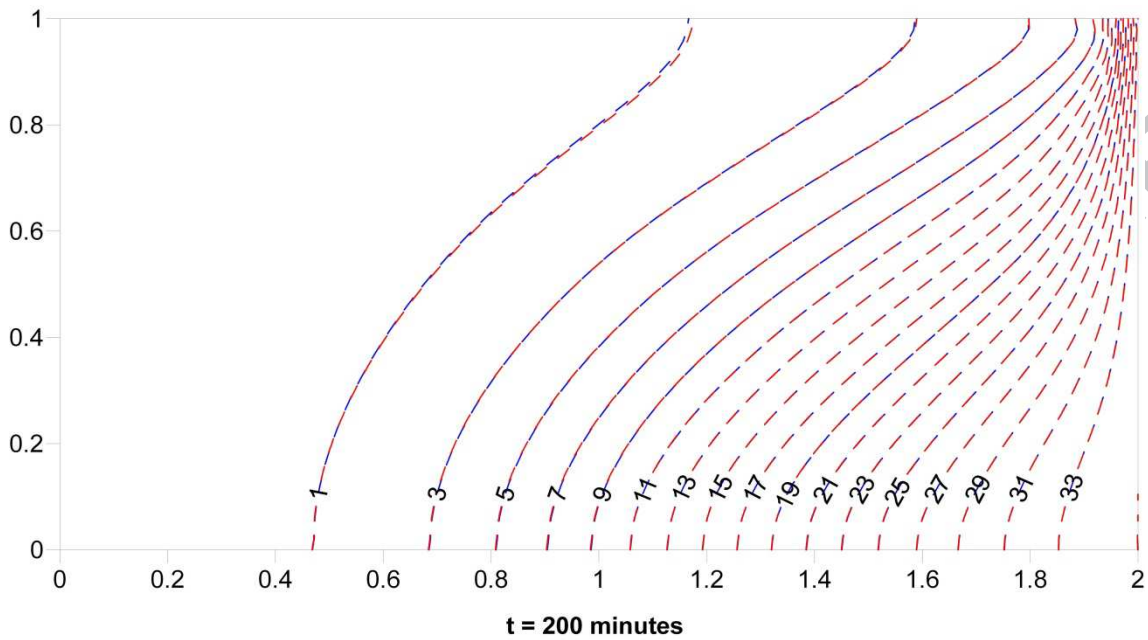


Figure 9



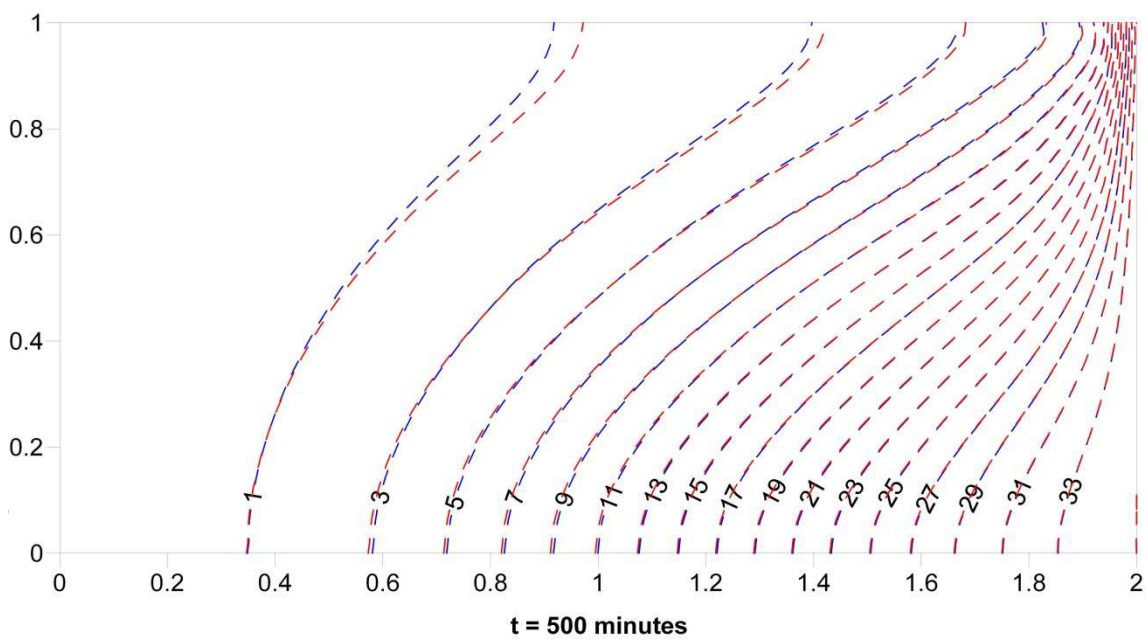


Figure 10

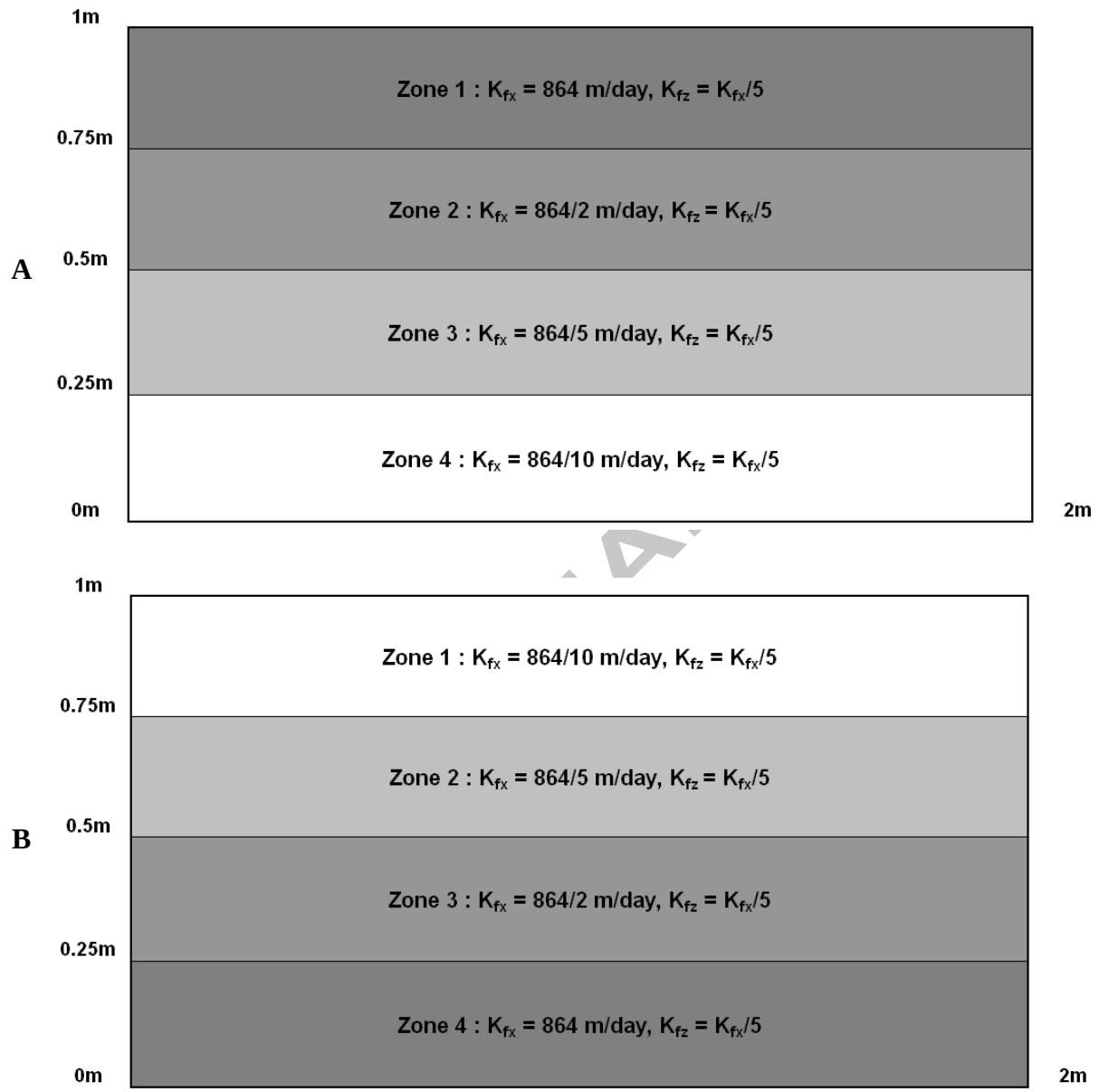


Figure 11

Figure 11

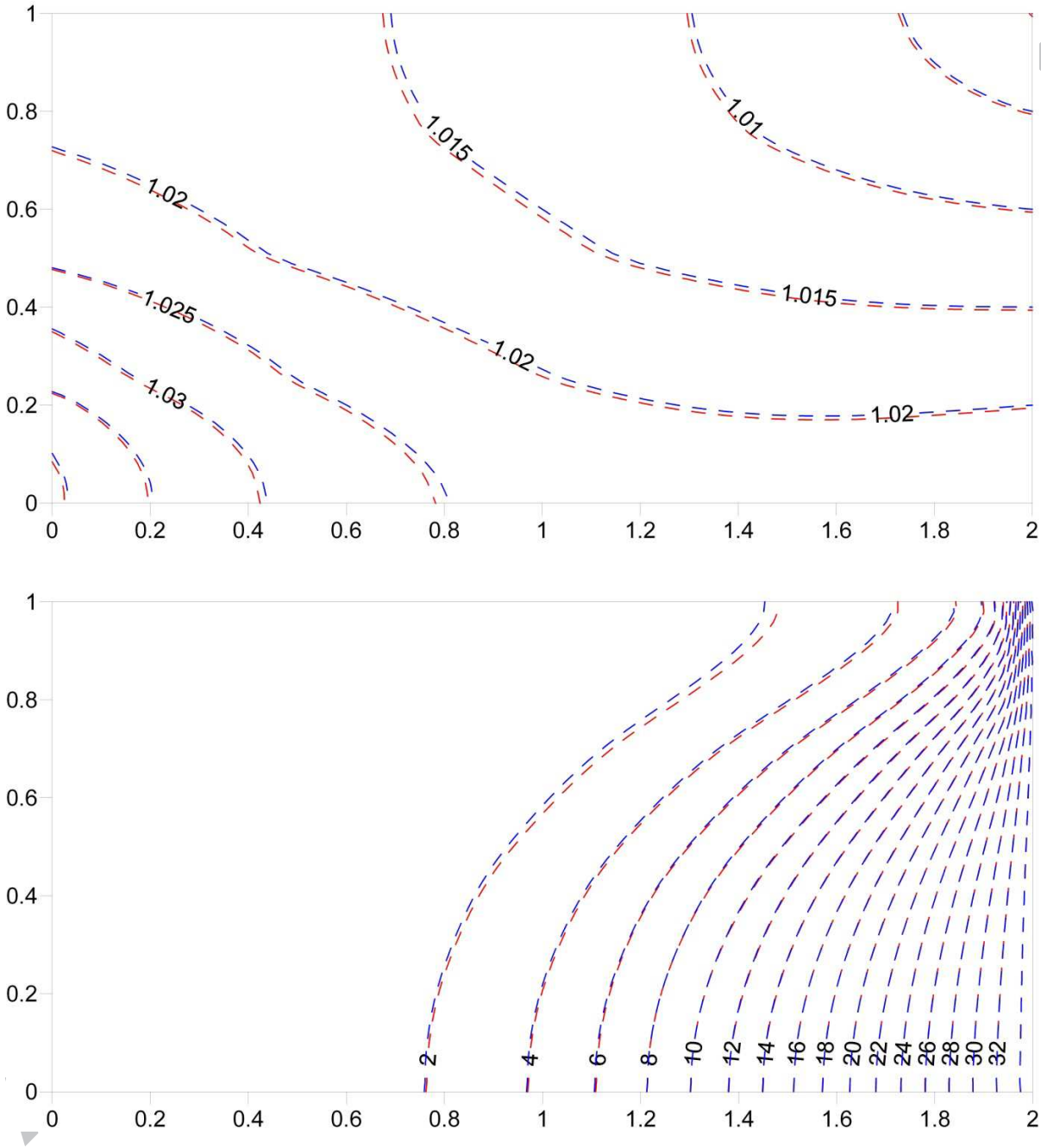


Figure 12

Figure 12

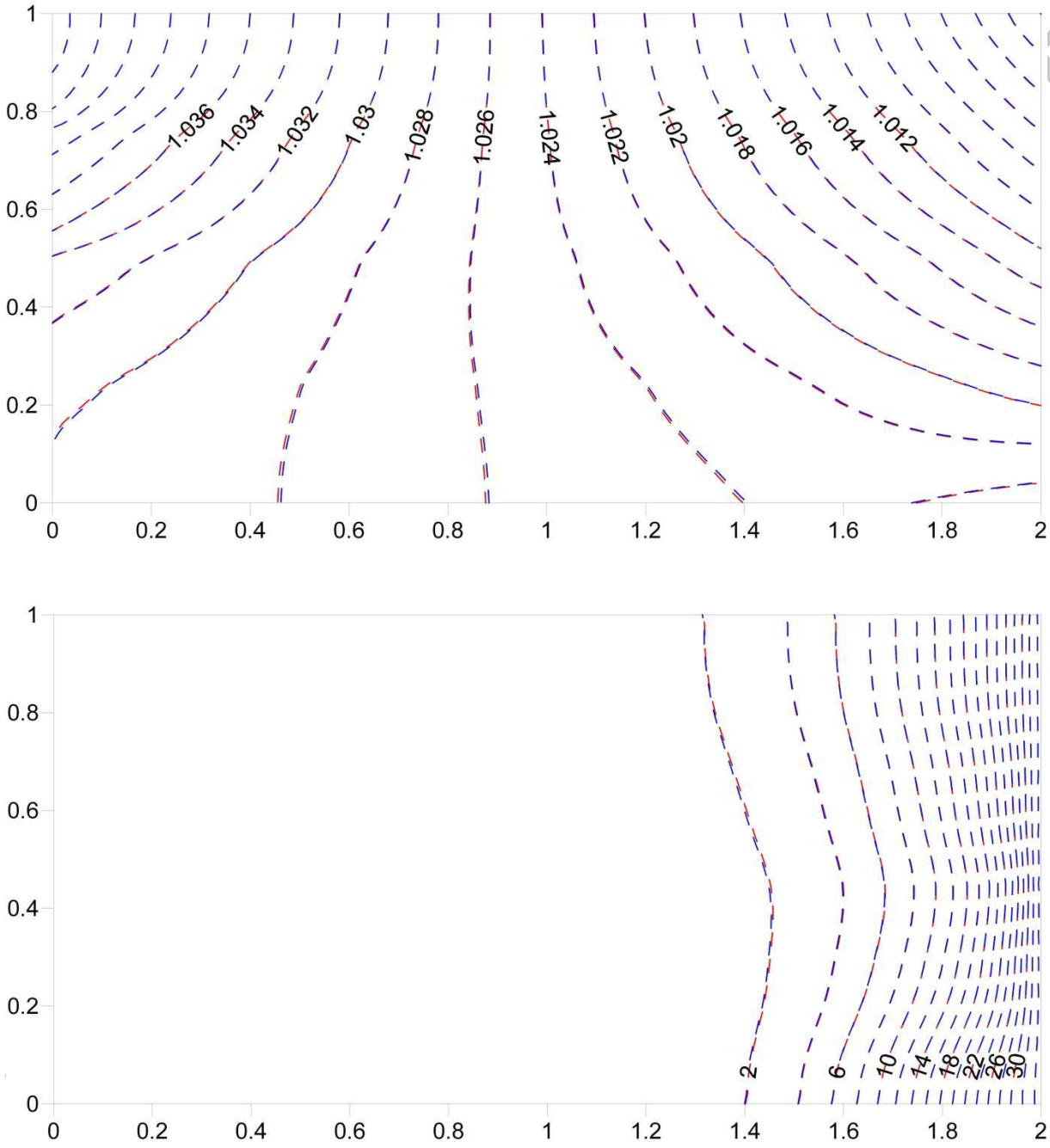


Figure 13

Figure 13

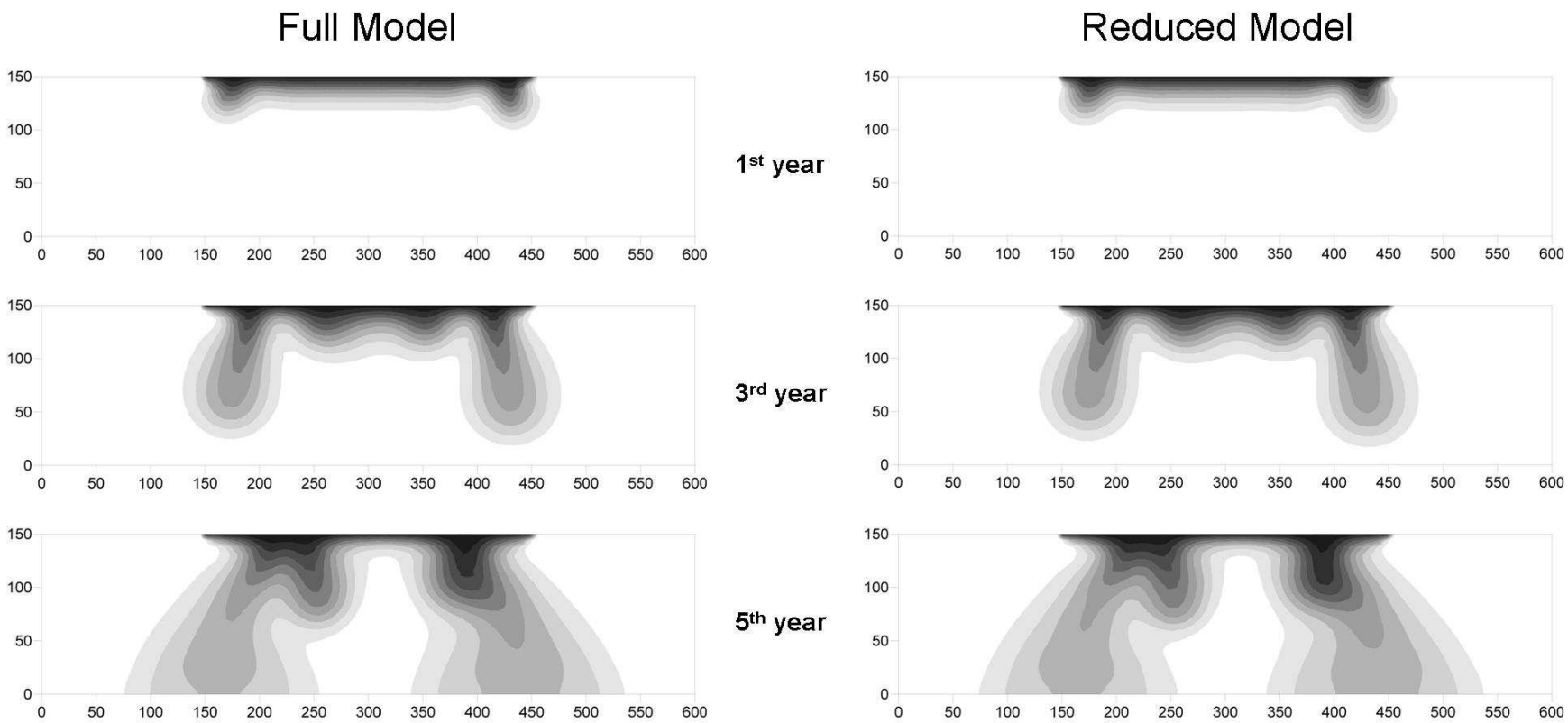


Figure 14

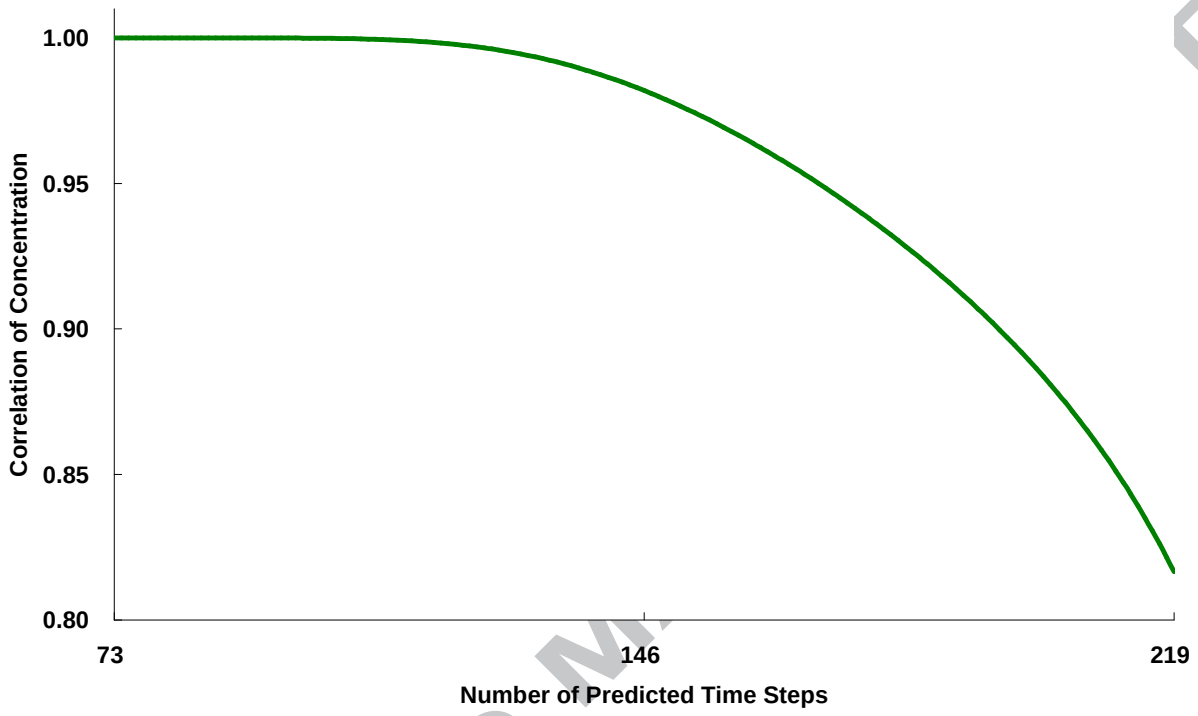
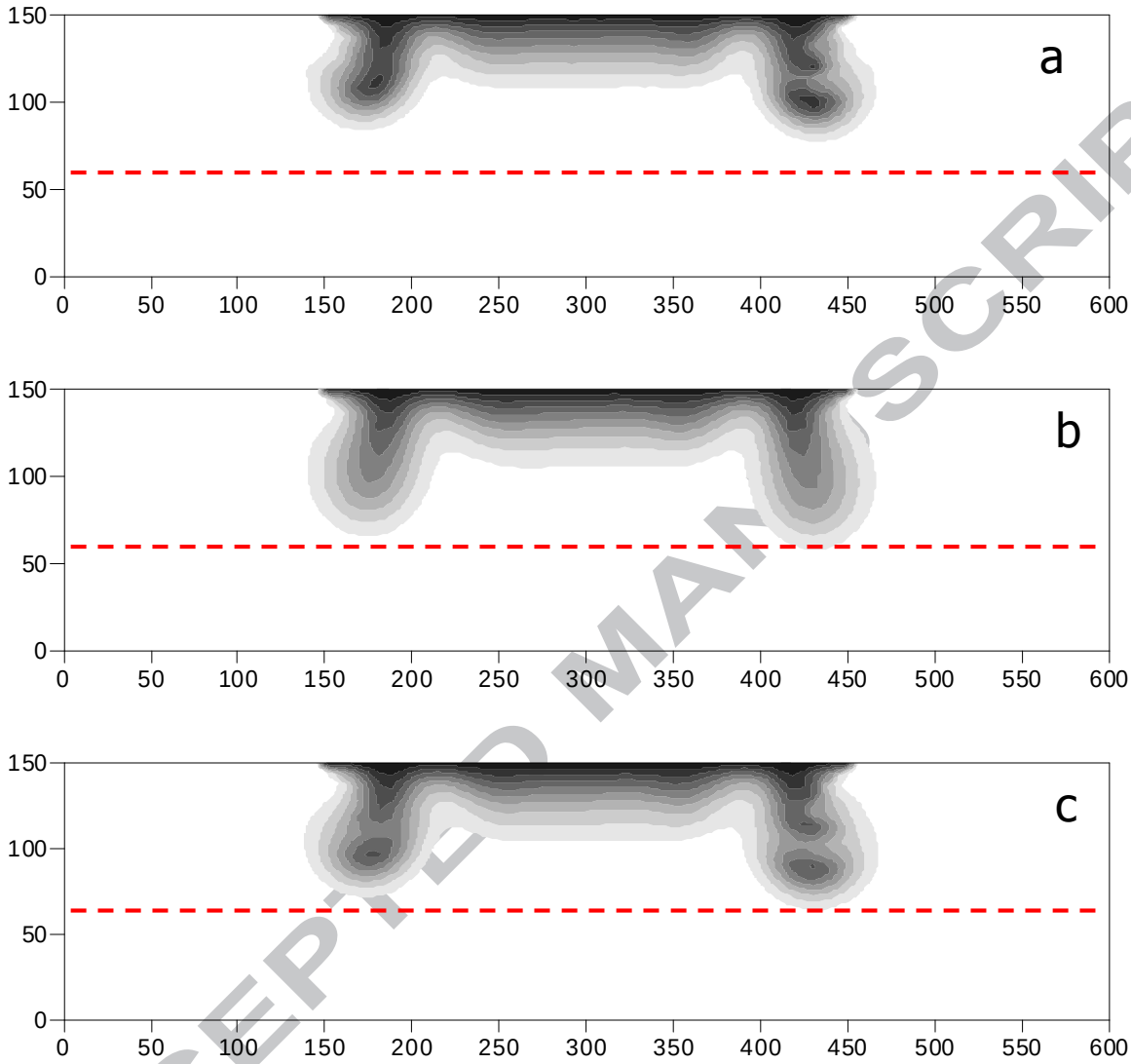


Figure 15

Figure 15



Highlights

1. Developing mathematical and numerical methods to simulate variable-density flow and solute transport;
2. A model reduction technique called Proper Orthogonal Decomposition designed for both linear and nonlinear models;
3. Model application to two classic variable-density flow and solute transport cases, the Henry problem and the Elder problem.
4. High efficiency of the developed method